



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00, 15/12, C07K 14/705, 14/435		A1	(11) International Publication Number: WO 99/37757 (43) International Publication Date: 29 July 1999 (29.07.99)
(21) International Application Number: PCT/US99/01259 (22) International Filing Date: 21 January 1999 (21.01.99)		(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority Data: 60/072,197 22 January 1998 (22.01.98) US		Published <i>With international search report.</i>	
(71) Applicants (for all designated States except US): ADMINISTRATORS OF THE TULANE EDUCATIONAL FUND [US/US]; 1430 Tulane Avenue, SL76, New Orleans, LA 70112-2699 (US). INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE [FR/FR]; 101, rue de Tolbiac, F-75654 Paris Cedex 13 (FR).			
(72) Inventors; and (75) Inventors/Applicants (for US only): HAMMOND, Timothy, G. [AU/US]; 1418 Bordeaux Street, New Orleans, LA 70115 (US). VEROUST, Pierre, J. [FR/FR]; Hôpital Tenon, INSERM U489, 4, rue de la Chine, F-75020 Paris (FR).			
(74) Agent: ADLER, Benjamin, A.; McGregor & Adler, 8011 Candle Lane, Houston, TX 77071 (US).			

(54) Title: CUBILIN PROTEIN, DNA SEQUENCES ENCODING CUBILIN AND USES THEREOF

(57) Abstract

The present invention provides novel renal receptors for ligands. Cubilin and megalin are representative examples of such renal receptors. Also provided are potential uses of these renal receptors for treating toxicity in various tissues and for detecting renal damage.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

CUBILIN PROTEIN, DNA SEQUENCES ENCODING CUBILIN AND USES THEREOF

5

BACKGROUND OF THE INVENTION

10 Cross-reference to Related Application

This patent application claims benefit of provisional patent application U.S. Serial number 60/072,197, filed January 22, 1998, now abandoned.

15 Federal Funding Notice

The present invention was funded in part by NIH grant DK46117. Consequently, the United States government has certain rights in this invention.

20 Field of the Invention

The present invention relates generally to the fields of molecular biology, biochemistry and medical therapy. More specifically, the present invention relates to major renal receptors for low molecular weight proteins and potential uses of the receptors for therapy to prevent renal toxicity.

Description of the Related Art

Countless proteins of small or intermediate molecular weight, filtered freely or partially through the renal glomerulus,

are bound by scavenger pathway receptors on the luminal surface of proximal tubular cells for reuptake (Batuman et al., 1990; Birn et al., 1997; Christensen et al., 1995; Saito et al., 1994). These proteins are then transcytosed back into the circulation, or 5 degraded, releasing amino acids for fresh protein synthesis. The scavenger pathway receptors of the proximal tubular are an essential physiological defense against the urinary loss of a diverse array of plasma proteins essential to homeostatic functions from coagulation to lipid metabolism. Unfortunately, 10 exposure of the scavenger receptors to unusually high concentrations of ligands due to overproduction of a ligand such as myeloma light chains, introduction of a freely filtered drug such as gentamicin, or increased glomerular permeability can disrupt the physiological balance, resulting in severe nephrotoxicity.

15 The relative contribution of proximal and distal elements to the development of protein nephrotoxicity remains controversial and ill-defined. For light chain nephrotoxicity if they fail to be reabsorbed proximally, the ligands are delivered into the distal tubular segments of the nephron, where they 20 precipitate as casts in combination with Tamm-Horsfall protein (Huang et al., 1997; Weiss et al., 1981; Winearls, 1995) .the severity of the renal dysfunctional correlates with the degree of (distal) cast formation (Myatt, 1994; Winearls, 1995). However, some light chains are associated with a pure (proximal) Fanconi 25 syndrome. Myoglobin on the other hand is associated with little (distal) cast formation, but marked proximal tubular damage, with clinical acute tubular necrosis (Paller, 1988; Zager, 1991).

Immunoglobulin light chains are filtered at the glomerulus and endocytosed in the proximal tubule (Batuman et

al., 1990; Batuman et al., 1997). In overproduction states, such as multiple myeloma, light chains, also known as Bence-Jones proteins, may produce nephrotoxicity. It was shown previously that free κ -and λ -light chain isotypes bind to a single class of renal 5 proximal tubular receptors which facilitate internalization and degradation (Batuman et al., 1997). To date, however, the receptor(s) which mediate endocytosis of light chains in the proximal tubule have not been characterized.

It has long been postulated that glycoproteins 10 expressed at the apical pole of proximal tubule cells of the kidney acted as scavenger pathway receptors. The only known and cloned receptor until now is megalin, a fairly abundant proximal tubule protein, also known as gp330 or the "Heymann antigen". Megalin is a classic single transmembrane domain giant 15 glycoprotein receptor (Saito et al., 1994), which belongs to the LDLR family (Yamamoto, 1984), and is closely related to the α_2 -macroglobulin receptor, which is not expressed in the kidney (Moestrup, 1994). Characterization of megalin revealed that, like the α_2 M receptor, it was a multiligand receptor. Of particular 20 interest for the renal pathology, megalin binds tPA and urokinase in complex with the corresponding inhibitor, but is also a polybasic drug receptor, binding ligands such as the aminoglycoside antibiotics (Moestrup et al., 1995).

Thus the prior art is deficient in the lack of renal 25 receptors for toxic, physiological, and pathological proteins and drugs (such as myeloma light chains) and more generally, components that may gain access to the proximal tubule fluid. Further, the prior art is deficient in the lack of effective means of preventing renal toxicity by utilizing renal binding proteins or

fragments thereof for such components. The present invention fulfills these long-standing needs and desires in the art.

SUMMARY OF THE INVENTION

5

The present invention discloses renal binding proteins for ligands. Also disclosed are the potential uses of these proteins for therapy to prevent renal toxicity or other types of toxicity.

In one embodiment of the present invention, there is
10 provided a DNA encoding a cubilin protein selected from the group consisting of: (a) isolated DNA which encodes a cubilin protein; (b) isolated DNA which hybridizes to isolated DNA of (a) and which encodes a cubilin protein; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) in codon sequence due to the
15 degeneracy of the genetic code, and which encodes a cubilin protein. Preferably, the DNA has the sequence shown in SEQ ID No. 1, and cubilin protein has the amino acid sequence shown in SEQ ID No. 2. Still preferably, the DNA is expressed in the tissues like kidney, spleen, brain, liver, heart and thyroid.

20 In one embodiment of the present invention, there is a vector capable of expressing the DNA adapted for expression in a recombinant cell and regulatory elements necessary for expression of the DNA in the cell. Specifically, the DNA encodes a cubilin protein.

25 In another embodiment of the present invention, there is a host cell transfected with the vector expressing a cubilin protein. Specifically, the host cell is selected from the group consisting of bacterial cells, mammalian cells, plant cells and insect cells. More specifically, the bacterial cell is *E. coli*.

In another embodiment of the present invention, there is provided isolated and purified cubilin protein or fragment coded for by DNA selected from the group consisting of: (a) isolated DNA which encodes a cubilin protein or fragment; (b) 5 isolated DNA which hybridizes to isolated DNA of (a) and which encodes a cubilin protein or fragment; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) in codon sequence due to the degeneracy of the genetic code, and which encodes a cubilin protein or fragment. Preferably, the protein has the amino 10 acid sequence shown in SEQ ID No. 2, and the fragment has amino acid sequence consisting of one or more of the sequences selected from the group consisting of SEQ ID Nos. 21-27.

In another embodiment of the present invention, there is provided a method of detecting expression of the cubilin protein or fragment in a sample, comprising the steps of: (a) 15 contacting mRNA obtained from the sample with a labeled hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

In another preferred embodiment of the present 20 invention, there is provided a pharmaceutical composition comprising the cubilin protein or fragment and a pharmaceutically acceptable carrier. Such composition can be used for treating or reducing nephrotoxicity or other types of toxicity in an in-need individual.

In still another embodiment of the present invention, 25 there is provided a receptor for a variety of ligands, comprising a cluster of EGF repeats and a cluster of CUB domains. Specifically, the receptor is cubilin and the ligand is selected from the group consisting of immunoglobulin light chain, myoglobin, intrinsic

factor-vitamin B₁₂, metallothionein, β -2-microglobulin, amyloid, hemoglobin, haptoglobin, interferon, insulin, cytochrome c, lysozyme, transferrin, transthyretin, polybasic drugs, apolipoprotein AI, high density lipoprotein and receptor related protein. More specifically, a representative example of polybasic drug is gentamicin. Representative examples of immunoglobulin light chain include κ -light chain and λ -light chain.

In still yet another embodiment of the present invention, there is provided a method of detecting renal damage by measuring the level of cubilin in the urine of an individual suspected to have such damage. If the urinary cubilin level is lower than that of a normal individual, the test individual might have chronic renal damage. On the other hand, the test individual might have renal damage of acute origin if the urinary cubilin level is higher than that of a normal individual.

Other and further aspects, features, and advantages of the present invention will be apparent from the following description of the presently preferred embodiments of the invention given for the purpose of disclosure.

20

BRIEF DESCRIPTION OF THE DRAWINGS

So that the matter in which the above-recited features, advantages and objects of the invention, as well as others which will become clear, are attained and can be understood in detail, more particular descriptions of the invention briefly summarized above may be had by reference to certain embodiments thereof which are illustrated in the appended drawings. These drawings

form a part of the specification. It is to be noted, however, that the appended drawings illustrate preferred embodiments of the invention and therefore are not to be considered limiting in their scope.

5 **Figure 1** shows cDNA cloning of the rat yolk sac cubilin. The 11.6 kb cDNA sequence was encoded by three clones (b-d) from a rat yolk sac cell cDNA library. The ends of the cDNA were confirmed by sequencing 3'- and 5'-RACE products from a rat kidney cDNA library with end-ligated adaptors. A 10 polyadenylation signal followed by a poly(A) tail is present 0.4 kb downstream the open reading frame (0.4-11.2 kb). The position of the initially identified clone is indicated (a).

Figure 2 shows a Northern blot of yolk sac RNA using a cubilin RNA probe.

15 **Figure 3** shows amino acid sequence of rat cubilin (SEQ ID NO. 2) as deduced by cDNA cloning. The predicted 20 amino acid endoplasmatic import signal sequence is shown in italic type. Potential glycosylation sites are indicated by asterisks. The sequence verified by protein sequencing of tryptic peptides are 20 shown in boldface type. The estimated mass of the peptide backbone is 396,953 kDa and pI is 5.6.

Figure 4 shows deglycosylation of cubilin purified by IF-B₁₂ affinity chromatography of rabbit renal cortex membranes. Reducing SDS-PAGE shows a reduction from 460 kDa (left lane) to 25 approximately 400 kDa (right lane) after treatment with peptide N-glycosidase F (PNGase F).

Figure 5 shows the extracellular modules of cubilin. **Figure 5A** shows schematic representation of the 460 kDa receptor (designated cubilin) and related developmental control

proteins, human bone morphogenic protein-1 (BMP-1), human tumor necrosis factor stimulating gene 6 (TSG-6), pig spermadhesin aqn3, and the *Drosophila* protein tollloid. The EGF repeats and CUB domains encode the whole protein except the 110 residues after the signal peptide. **Figure 5B** shows a dot plot display of the high internal homology of the CUB domains in cubilin.

Figure 6A shows alignment of the EGF repeats in cubilin and homologous repeats in CUB domain-containing proteins (Bmp-1, tollloid protein, C1s) and in human fibrillin-1 (SEQ ID Nos. 6-20). The consensus residues for calcium binding are indicated at the bottom of the figure. **Figure 6B** shows alignment of cubilin CUB domains and CUB domains in the development control proteins shown in Figure 5, *Xenopus laevis* Uvs-2 and human C1s (SEQ ID Nos. 21-36).

Figure 7 shows release of membrane-associated cubilin by non-enzymatic and non-solubilizing procedures. Western blot analysis with anti-cubilin antibody (top panel) and anti-megalin antibody (bottom panel). Lane 1: yolk sac BN cells; lane 2: ileal mucosa; lane 3: membrane phase after mechanical grinding of renal cortex; lane 4: fluid phase after mechanical grinding of renal cortex membranes, lanes 5-11: fluid phase of renal cortex membranes after a 1-h incubation in PBS (lane 5), PBS, phosphorylethanolamine, and heparin (lane 6), PBS, phosphorylethanolamine, heparine, and EDTA (lane 7), PBS and heparin (lane 8), PBS and EDTA (lane 9), PBS, heparin, and EDTA (lane 10), PBS and phosphorylethanolamine (lane 11).

Figure 8 shows immunocytochemical localization of megalin (10 nm gold) and cubilin (5 nm gold) in the apical part of

rat renal proximal tubule cell (**Figure 8A**) and epithelial cell of rat yolk sac (**Figure 8B**). The two proteins are colocalized in apical endocytic invaginations (AI), small (SE) and large (LE) endosomes, the small gold-particles (cubilin) being indicated by 5 large arrowheads. Colocalization is also seen in dense apical tubules (cubilin, small arrowheads). Microvilli (MV) of the proximal tubule are labeled for both proteins, whereas very little labeling is seen on yolk sac microvilli. Bars, 0.1 μ m.

Figure 9 shows binding of cubilin to megalin as 10 determined by affinity chromatography. 125 I-cubilin was applied to a megalin-Sepharose-4B column (•) or a blank Sepharose-4B column (o). The inset demonstrates autoradiography of SDS-PAGE of 125 I-cubilin and the eluted fraction 13 (#13). Bound radioactivity was eluted by the addition (arrow) of 10 mM EDTA 15 to the running buffer.

Figure 10 shows characterization of the cubilin-megalin interaction by surface plasmon resonance analysis. Rabbit megalin was immobilized to a sensor chip and the on rates and off rates for the binding of cubilin was recorded by flow 20 of 20 nM purified cubilin along the chip surface. For control, cubilin was subjected to a blank chip. The values displayed are the recordings from the megalin-chip subtracted from the recordings from the blank chip. **Figure 10A** shows sensorgram of the binding of cubilin to megalin. The binding curves in the 25 presence of 10 mM EDTA or after prebinding of RAP to megalin are also shown. **Figure 10B** demonstrates the formation of an IF-B₁₂-cubilin megalin complex by subsequent flow with cubilin, running buffer and IF-B₁₂. Evaluation of the binding data suggests

a complex binding. By fitting the binding data to a one-binding-site model a K_d of 7 nM was measured.

Figure 11 shows multiple lines of evidence suggesting that cubilin binds light chains: **Figure 11A** shows 5 Western blot analysis with a holo-gp280 polyclonal antiserum (left gel) and SDS-PAGE result (right gel). Cubilin at 460 kDa is accompanied by a prominent band at the region of 56 kDa. **Figure 11B** shows coomassie-stained two-dimensional gel with pH gradient from 4-8 on *abscissa* and molecular weight on the 10 *ordinate* which demonstrates relative protein abundance.

Figure 12 shows direct binding analysis of cubilin and myeloma light chains by surface plasmon resonance. **Figure 12A** shows binding of cubilin to immobilized κ -light chains is dose dependent with rapid low affinity association and dissociation 15 kinetics. **Figure 12B** shows competition experiment which further demonstrates the specificity of binding of cubilin to immobilized κ -light chains. A sample of cubilin (100 nM) was incubated with κ -light chains (10 or 490 μ M), or κ -light chains (10 or 490 μ M) prior to injecting the sample over the κ -light chain 20 surface. The binding of cubilin to the immobilized surfaces was reduced in the presence of κ -or λ -light chains in a dose-response fashion. Data are representative of experiments with 4 light chains on 3 chips. **Figure 12C** shows effect of temperature on 25 binding of cubilin (110 nM) to λ -light chains surface. LC, light chains: RU, response units.

Figure 13 shows displacement of light chain binding to rat renal cortical brush-border membranes by anti-cubilin and megalin antisera. **Figure 13A** shows that anti-cubilin polyclonal

antiserum (•) inhibits ^{125}I -labeled λ -light chain binding to rat renal brush-border membrane vesicles, but megalin antiserum (o) had no effect. **Figure 13B** shows vesicle-by-vesicle analysis of FITC-light chain binding by flow cytometry. Each panel depicts 5 2,000 vesicles, and each dot represents one vesicle. FITC-fluorescence on the *abscissa* is displayed against vesicle size on the *ordinate*. Representative of $n=8$. Note most but not all vesicles bind FITC-light chains in the left panel, and anti-cubilin antiserum reduces binding (right panel).

10 **Figure 14** shows effects of anti-cubilin antibody on light chain endocytosis. **Figure 14A** shows FITC- κ -light chain endocytosis by yolk sac epithelial cells (BN/MSV) at 30 min. **Figure 14B** shows time course of the endocytosis of FITC- λ -light chain in BN/MSV cells in the presence and absence of anti-cubilin 15 antibody over 40 min. LC, light chain; Ab, antibody.

20 **Figure 15** shows direct effects of light chains on endosomal fusion reconstituted *in vitro*. The λ -light chains were loaded into rat renal cortical intermicrovillar cleft at 400 μM by addition to the homogenization buffer. Fusion reconstituted *in* **vitro** in light chain-loaded membranes was inhibited compared with albumin-loaded control membranes. Values are mean \pm standard error for $n=8$, $p < 0.05$ by unpaired *t*-test.

25 **Figure 16** shows fluorescein-light chain binding to renal brush-border membrane vehicles in the presence of gentamicin by flow cytometry, demonstrating megalin is a light chain receptor.

Figure 17 shows extra renal expression of cubilin. **Figure 17A** shows RT-PCR with 2 sets of primers. Lanes 1, 3, 5, 7

and 9 used a primer (tgcctaccacagccaaatga, SEQ ID No. 37) located in one of the 3' CUB domains; and lanes 2, 4, 6, 8 and 10 used another primer (agagccacaatgactgcag, SEQ ID No. 38) located in the end of the EGF regions. Lanes 1-2: RNA from spleen; lanes 3-4: 5 RNA from brain; lanes 5-6: RNA from liver; lanes 7-8: RNA from heart; lanes 9-10: RNA from Brown Norway (BN) rat immortalized yolk sac epithelial cells; lane 11: RNA from immortalized opossum kidney (OK) cells with megalin primers (SEQ ID Nos. 39-40). Figure 17B shows RT-PCR with cubilin primers (SEQ ID Nos. 37-10 38). Lane 1: kidney; lane 2: thyroid.

Figure 18 shows urine cubilin. Normal rat urine (lane 1) and 2 separate preparations of rat renal brush border (lanes 2-3) were separated by PAGE-SDS, transferred onto nitrocellulose and probed with anti-cubilin antibodies.

15

DETAILED DESCRIPTION OF THE INVENTION

The present invention discloses the molecular characterization of the 460 kDa epithelial glycoprotein that 20 functions as the receptor facilitating uptake of intrinsic factor-vitamin B₁₂ complexes in the intestine and kidney. The 3603 amino acid rat sequence has one cluster of 8 EGF-type domains followed by a cluster of 27 CUB domains accounting for 88% of the protein mass. The receptor, cubilin, has no similarity to known 25 endocytic receptors. Instead, it displays homology to EGF and CUB domain-containing proteins involved in fetal development. Cubilin is a peripheral membrane protein which can be released from renal cortex membranes by non-enzymatic and non-solubilizing procedures. Electron microscopic immuno-gold

labeling of rat yolk sac and renal proximal tubules revealed that the endocytic receptor megalin and cubilin strictly colocalize in the endocytic apparatus. Megalin-affinity chromatography and surface plasmon analysis demonstrated a calcium-dependent high affinity binding of cubilin to the extracellular part of megalin which thereby may assist the intracellular trafficking of this novel type of receptor.

Myeloma light chains are known to undergo receptor-mediated endocytosis in the kidney, however, the molecular identity of the receptor has not been characterized. The present studies provide several lines of evidence to identify cubilin (gp280), a giant glycoprotein receptor, which is preferentially expressed in endocytic scavenger pathways and which has potent effects on endosomal trafficking, as an endocytic receptor for immunoglobulin light chains. Binding showed dose and time-dependent saturability with low-affinity, high-capacity equilibrium binding parameters. The data demonstrate that cubilin plays a role in the endocytosis and trafficking of light chains in renal proximal tubule cells.

Agents that inhibit binding of light chains are ligands of cubilin. More generally, the proteins present in the urine of patients or dogs deficient in cubilin contain a variety of proteins including albumin which constitute ligands.

Independent evidence also suggests that light chains are ligands for megalin. These studies are important, both to understand the complex interactions of toxic and physiological ligands on proximal tubule scavenger pathway receptors, as well as the eventual development of clinical protective agents for

nephrotoxic damage mediated by ligands for cubilin and/or megalin.

If appearing herein, the following terms shall have the definitions set out below.

5 As used herein, the term "cDNA" shall refer to the DNA copy of the mRNA transcript of a gene.

As used herein, the term "derived amino acid sequence" shall mean the amino acid sequence determined by reading the triplet sequence of nucleotide bases in the cDNA.

10 As used herein the term "screening a library" shall refer to the process of using a labeled probe to check whether, under the appropriate conditions, there is a sequence complementary to the probe present in a particular DNA library. In addition, "screening a library" could be performed by PCR.

15 As used herein, the term "PCR" refers to the polymerase chain reaction that is the subject of U.S. Patent Nos. 4,683,195 and 4,683,202 to Mullis, as well as improvements now known in the art.

20 In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual" (1982); "DNA Cloning: A Practical Approach," Volumes I and II (D.N. Glover ed. 1985); "Oligonucleotide Synthesis" (M.J. Gait ed. 1984); "Nucleic Acid Hybridization" [B.D. Hames & S.J. Higgins eds. (1985)]; "Transcription and Translation" [B.D. Hames & S.J. Higgins eds. (1984)]; "Animal Cell Culture" [R.I. Freshney, ed. (1986)];

"Immobilized Cells And Enzymes" [IRL Press, (1986)]; B. Perbal, "A Practical Guide To Molecular Cloning" (1984).

A "vector" is a replicon, such as plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

A "DNA molecule" refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, or cytosine) in its either single stranded form, or a double-stranded helix. This term refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear DNA molecules (e.g., restriction fragments), viruses, plasmids, and chromosomes. In discussing the structure herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (i.e., the strand having a sequence homologous to the mRNA).

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences. A polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers,

polyadenylation signals, terminators, and the like, that provide for the expression of a coding sequence in a host cell.

An "expression control sequence" is a DNA sequence that controls and regulates the transcription and translation of 5 another DNA sequence. A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then translated into the protein encoded by the coding sequence.

10 A "signal sequence" can be included near the coding sequence. This sequence encodes a signal peptide, N-terminal to the polypeptide, that communicates to the host cell to direct the polypeptide to the cell surface or secrete the polypeptide into the media, and this signal peptide is clipped off by the host cell before 15 the protein leaves the cell. Signal sequences can be found associated with a variety of proteins native to prokaryotes and eukaryotes.

The term "oligonucleotide", as used herein in referring to the probe of the present invention, is defined as a molecule 20 comprised of two or more ribonucleotides, preferably more than three. Its exact size will depend upon many factors which, in turn, depend upon the ultimate function and use of the oligonucleotide.

The term "primer" as used herein refers to an oligonucleotide, whether occurring naturally as in a purified 25 restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product, which is complementary to a nucleic acid strand, is induced, i.e., in the presence of nucleotides and an inducing agent such as a DNA

polymerase and at a suitable temperature and pH. The primer may be either single-stranded or double-stranded and must be sufficiently long to prime the synthesis of the desired extension product in the presence of the inducing agent. The exact length of 5 the primer will depend upon many factors, including temperature, source of primer and use the method. For example, for diagnostic applications, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15-25 or more nucleotides, although it may contain fewer nucleotides.

10 The primers herein are selected to be "substantially" complementary to different strands of a particular target DNA sequence. This means that the primers must be sufficiently complementary to hybridize with their respective strands. Therefore, the primer sequence need not reflect the exact 15 sequence of the template. For example, a non-complementary nucleotide fragment may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided that the 20 primer sequence has sufficient complementarity with the sequence or hybridize therewith and thereby form the template for the synthesis of the extension product.

A cell has been "transformed" by exogenous or heterologous DNA when such DNA has been introduced inside the 25 cell. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transforming DNA may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transformed cell is one in

which the transforming DNA has become integrated into a chromosome so that it is inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA. A "clone" is a population of cells derived from a single cell or ancestor by mitosis. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

Two DNA sequences are "substantially homologous" when at least about 75% (preferably at least about 80%, and most preferably at least about 90% or 95%) of the nucleotides match over the defined length of the DNA sequences. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Maniatis et al., *supra*; DNA Cloning, Vols. I & II, *supra*; Nucleic Acid Hybridization, *supra*.

A "heterologous" region of the DNA construct is an identifiable segment of DNA within a larger DNA molecule that is not found in association with the larger molecule in nature. Thus, when the heterologous region encodes a mammalian gene, the gene will usually be flanked by DNA that does not flank the mammalian genomic DNA in the genome of the source organism. In another example, coding sequence is a construct where the coding sequence itself is not found in nature (e.g., a cDNA where the genomic coding sequence contains introns, or synthetic

sequences having codons different than the native gene). Allelic variations or naturally-occurring mutational events do not give rise to a heterologous region of DNA as defined herein.

The labels most commonly employed for these studies 5 are radioactive elements, enzymes, chemicals which fluoresce when exposed to ultraviolet light, and others. A number of fluorescent materials are known and can be utilized as labels. These include, for example, fluorescein, rhodamine, auramine, Texas Red, AMCA blue and Lucifer Yellow. A particular detecting 10 material is anti-rabbit antibody prepared in goats and conjugated with fluorescein through an isothiocyanate.

Proteins can also be labeled with a radioactive element or with an enzyme. The radioactive label can be detected by any of the currently available counting procedures. The preferred 15 isotope may be selected from ^3H , ^{14}C , ^{32}P , ^{35}S , ^{36}Cl , ^{51}Cr , ^{57}Co , ^{58}Co , ^{59}Fe , ^{90}Y , ^{125}I , ^{131}I , and ^{186}Re .

Enzyme labels are likewise useful, and can be detected by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or gasometric techniques. 20 The enzyme is conjugated to the selected particle by reaction with bridging molecules such as carbodiimides, diisocyanates, glutaraldehyde and the like. Many enzymes which can be used in these procedures are known and can be utilized. The preferred are peroxidase, β -glucuronidase, β -D-glucosidase, β -D- 25 galactosidase, urease, glucose oxidase plus peroxidase and alkaline phosphatase. U.S. Patent Nos. 3,654,090, 3,850,752, and 4,016,043 are referred to by way of example for their disclosure of alternate labeling material and methods.

A particular assay system developed and utilized in the art is known as a receptor assay. In a receptor assay, the material to be assayed is appropriately labeled and then certain cellular test colonies are inoculated with a quantity of both the label after which binding studies are conducted to determine the extent to which the labeled material binds to the cell receptors. In this way, differences in affinity between materials can be ascertained.

An assay useful in the art is known as a "cis/trans" assay. Briefly, this assay employs two genetic constructs, one of which is typically a plasmid that continually expresses a particular receptor of interest when transfected into an appropriate cell line, and the second of which is a plasmid that expresses a reporter such as luciferase, under the control of a receptor/ligand complex. Thus, for example, if it is desired to evaluate a compound as a ligand for a particular receptor, one of the plasmids would be a construct that results in expression of the receptor in the chosen cell line, while the second plasmid would possess a promoter linked to the luciferase gene in which the response element to the particular receptor is inserted. If the compound under test is an agonist for the receptor, the ligand will complex with the receptor, and the resulting complex will bind the response element and initiate transcription of the luciferase gene. The resulting chemiluminescence is then measured photometrically, and dose response curves are obtained and compared to those of known ligands. The foregoing protocol is described in detail in U.S. Patent No. 4,981,784.

As used herein, the term "host" is meant to include not only prokaryotes but also eukaryotes such as yeast, plant and

animal cells. A recombinant DNA molecule or gene which encodes cubilin protein of the present invention can be used to transform a host using any of the techniques commonly known to those of ordinary skill in the art. Especially preferred is the use of a 5 vector containing coding sequences for the gene which encodes cubilin protein of the present invention for purposes of prokaryote transformation.

Prokaryotic hosts may include *E. coli*, *S. typhimurium*, *Serratia marcescens* and *Bacillus subtilis*. Eukaryotic hosts include 10 yeasts such as *Pichia pastoris*, mammalian cells and insect cells.

In general, expression vectors containing promoter sequences which facilitate the efficient transcription of the inserted DNA fragment are used in connection with the host. The expression vector typically contains an origin of replication, 15 promoter(s), terminator(s), as well as specific genes which are capable of providing phenotypic selection in transformed cells. The transformed hosts can be fermented and cultured according to means known in the art to achieve optimal cell growth.

The invention includes a substantially pure DNA 20 encoding a cubilin protein, a strand of which DNA will hybridize at high stringency to a probe containing a sequence of at least 15 consecutive nucleotides of SEQ ID NO:1. The protein encoded by the DNA of this invention may share at least 80% sequence identity (preferably 85%, more preferably 90%, and most 25 preferably 95%) with the amino acids listed in SEQ ID NO. 2. More preferably, the DNA includes the coding sequence of the nucleotides of SEQ ID NO:1, or a degenerate variant of such a sequence.

The probe to which the DNA of the invention hybridizes preferably consists of a sequence of at least 20 consecutive nucleotides, more preferably 40 nucleotides, even more preferably 50 nucleotides, and most preferably 100 5 nucleotides or more (up to 100%) of the coding sequence of the nucleotides listed in SEQ ID NO. 1 or the complement thereof. Such a probe is useful for detecting expression of cubilin in a human cell by a method such as a method including the steps of (a) contacting mRNA obtained from the cell with the labeled 10 hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

This invention also includes a substantially pure DNA containing a sequence of at least 15 consecutive nucleotides (preferably 20, more preferably 30, even more preferably 50, 15 and most preferably all) of the nucleotides listed in SEQ ID NO. 1.

By "high stringency" is meant DNA hybridization and wash conditions characterized by high temperature and low salt concentration, e.g., wash conditions of 65°C at a salt concentration of approximately 0.1 x SSC, or the functional equivalent thereof. 20 For example, high stringency conditions may include hybridization at about 42°C in the presence of about 50% formamide; a first wash at about 65°C with about 2 x SSC containing 1% SDS; followed by a second wash at about 65°C with about 0.1 x SSC.

25 By "substantially pure DNA" is meant DNA that is not part of a milieu in which the DNA naturally occurs, by virtue of separation (partial or total purification) of some or all of the molecules of that milieu, or by virtue of alteration of sequences that flank the claimed DNA. The term therefore includes, for

example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment 5 produced by polymerase chain reaction (PCR) or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence, e.g., a fusion protein. Also included is a recombinant DNA which includes a portion of 10 the nucleotides listed in SEQ ID NO. 1 encoding an alternative splice variant of cubilin.

The DNA may have at least about 70% sequence identity to the coding sequence of the nucleotides listed in SEQ ID NO:1, preferably at least 75% (e.g. at least 80%); and most 15 preferably at least 90%. The identity between two sequences is a direct function of the number of matching or identical positions. When a subunit position in both of the two sequences is occupied by the same monomeric subunit, e.g., if a given position is occupied by an adenine in each of two DNA molecules, then they 20 are identical at that position. For example, if 7 positions in a sequence of 10 nucleotides in length are identical to the corresponding positions in a second 10-nucleotide sequence, then the two sequences have 70% sequence identity. The length of comparison sequences will generally be at least 50 nucleotides, 25 preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 100 nucleotides. Sequence identity is typically measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics

Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705).

The present invention comprises a vector comprising a DNA sequence coding for a human cubilin protein and said vector is capable of replication in a host which comprises, in operable linkage: a) an origin of replication; b) a promoter; and c) a DNA sequence coding for said protein. Preferably, the vector of the present invention contains a portion of the DNA sequence shown in SEQ ID No. 1. A "vector" may be defined as a replicable nucleic acid construct, e.g., a plasmid or viral nucleic acid. Vectors may be used to amplify and/or express nucleic acid encoding cubilin protein. An expression vector is a replicable construct in which a nucleic acid sequence encoding a polypeptide is operably linked to suitable control sequences capable of effecting expression of the polypeptide in a cell. The need for such control sequences will vary depending upon the cell selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter and/or enhancer, suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and translation. Methods which are well known to those skilled in the art can be used to construct expression vectors containing appropriate transcriptional and translational control signals. See for example, the techniques described in Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual* (2nd Ed.), Cold Spring Harbor Press, N.Y. A gene and its transcription control sequences are defined as being "operably linked" if the transcription control sequences effectively control the transcription of the gene. Vectors of the invention include, but are not limited to, plasmid vectors and

viral vectors. Preferred viral vectors of the invention are those derived from retroviruses, adenovirus, adeno-associated virus, SV40 virus, or herpes viruses.

By a "substantially pure protein" is meant a protein which has been separated from at least some of those components which naturally accompany it. Typically, the protein is substantially pure when it is at least 60%, by weight, free from the proteins and other naturally-occurring organic molecules with which it is naturally associated *in vivo*. Preferably, the purity of the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight. A substantially pure cubilin protein may be obtained, for example, by extraction from a natural source; by expression of a recombinant nucleic acid encoding an cubilin polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, e.g., column chromatography such as immunoaffinity chromatography using an antibody specific for cubilin, polyacrylamide gel electrophoresis, or HPLC analysis. A protein is substantially free of naturally associated components when it is separated from at least some of those contaminants which accompany it in its natural state. Thus, a protein which is chemically synthesized or produced in a cellular system different from the cell from which it naturally originates will be, by definition, substantially free from its naturally associated components. Accordingly, substantially pure proteins include eukaryotic proteins synthesized in *E. coli*, other prokaryotes, or any other organism in which they do not naturally occur.

In addition to substantially full-length proteins, the invention also includes fragments (e.g., antigenic fragments) of

the cubilin protein (SEQ ID No. 2). As used herein, "fragment," as applied to a polypeptide, will ordinarily be at least 10 residues, more typically at least 20 residues, and preferably at least 30 (e.g., 50) residues in length, but less than the entire, intact

5 sequence. Fragments of the cubilin protein can be generated by methods known to those skilled in the art, e.g., by enzymatic digestion of naturally occurring or recombinant cubilin protein, by recombinant DNA techniques using an expression vector that encodes a defined fragment of cubilin, or by chemical synthesis.

10 The ability of a candidate fragment to exhibit a characteristic of cubilin (e.g., binding to an antibody specific for cubilin, or binding to a known ligand of cubilin) can be assessed by methods described herein. Purified cubilin or antigenic fragments of cubilin can be used to generate new antibodies or to test existing

15 antibodies (e.g., as positive controls in a diagnostic assay) by employing standard protocols known to those skilled in the art. Included in this invention are polyclonal antisera generated by using cubilin or a fragment of cubilin as the immunogen in, e.g., rabbits. Standard protocols for monoclonal and polyclonal

20 antibody production known to those skilled in this art are employed. The monoclonal antibodies generated by this procedure can be screened for the ability to identify recombinant cubilin cDNA clones, and to distinguish them from known cDNA clones.

25 Further included in this invention are cubilin proteins or fragments which are encoded at least in part by portions of SEQ ID NO. 2, e.g., products of alternative mRNA splicing or alternative protein processing events, or in which a section of cubilin sequence has been deleted. The fragment, or the intact

cubilin polypeptide, may be covalently linked to another polypeptide, e.g. which acts as a label, a ligand or a means to increase antigenicity. The lack of cross inhibition of a number of ligands suggests that specific therapeutic components can be 5 produced.

Also within the invention is a method of detecting cubilin protein or fragment in a biological sample, which includes the steps of contacting the sample with the labelled antibody, e.g., radioactively tagged antibody specific for cubilin, and 10 determining whether the antibody binds to a component of the sample.

A standard Northern blot assay can be used to ascertain the relative amounts of cubilin mRNA in a cell or tissue obtained from a patient, in accordance with conventional 15 Northern hybridization techniques known to those persons of ordinary skill in the art. This Northern assay uses a hybridization probe, e.g. radiolabelled cubilin cDNA, either containing the full-length, single stranded DNA having a sequence complementary to SEQ ID NO. 1, or a fragment of that DNA sequence at least 20 (preferably at least 30, more preferably at least 50, and most 20 preferably at least 100 consecutive nucleotides in length). The DNA hybridization probe can be labelled by any of the many different methods known to those skilled in this art. RNA probes can also be similarly utilized.

25 The present invention is also directed to a vector capable of expressing the DNA of the present invention adapted for expression in a recombinant cell and regulatory elements necessary for expression of the DNA in the cell. Preferably, the

vector contains DNA encoding a cubilin protein having the amino acid sequence shown in SEQ ID No. 2.

The present invention is also directed to a host cell transfected with the vector described herein, said vector 5 expressing a cubilin protein. Representative host cells include consisting of bacterial cells, mammalian cells and insect cells.

The present invention is also directed to a isolated and purified cubilin protein or fragment coded for by DNA selected from the group consisting of: (a) isolated DNA which encodes a 10 cubilin protein or fragment; (b) isolated DNA which hybridizes to isolated DNA of (a) and which encodes a cubilin protein or fragment; and (c) isolated DNA differing from the isolated DNAs of (a) and (b) in codon sequence due to the degeneracy of the genetic code, and which encodes a cubilin protein or fragment. 15 Preferably, the isolated and purified cubilin protein has the amino acid sequence shown in SEQ ID No. 2, and the fragment has amino acid sequence consisting of one or more of the sequences selected from the group consisting of SEQ ID Nos. 21-27.

The present invention is also directed to a method of 20 detecting expression of the cubilin protein or fragment, comprising the steps of: (a) contacting mRNA obtained from a sample with a labeled hybridization probe; and (b) detecting hybridization of the probe with the mRNA.

The present invention is further directed to a 25 pharmaceutical composition comprising the cubilin protein or fragment and a pharmaceutically acceptable carrier. Such composition can be used for treating or reducing nephrotoxicity or other types of toxicity in an in-need individual.

In an additional embodiment, the present invention is directed to a receptor for a variety of ligands, comprising a cluster of EGF repeats and a cluster of CUB domains. Specifically, the receptor is cubilin and ligand selected from the group consisting of 5 immunoglobulin light chain, myoglobin, intrinsic factor-vitamin B₁₂, metallothionein, β-2-microglobulin, amyloid, hemoglobin, haptoglobin, interferon, insulin, cytochrome c, lysozyme, transferrin, transthyretin, polybasic drugs, low density lipoprotein, high density lipoprotein and receptor related protein. 10 A representative example of a polybasic drug is gentamicin. Representative examples of immunoglobulin light chain include κ-light chain and λ-light chain.

In still yet another embodiment of the present invention, there is provided a method of detecting renal damage 15 by measuring the level of cubilin in the urine of an individual suspected to have such damage. If the urinary cubilin level is lower than that of a normal individual, the test individual might have chronic renal damage, on the other hand, the test individual might have renal damage of acute origin if the urinary cubilin 20 level is higher than that of a normal individual.

The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to limit the present invention in any fashion.

25

EXAMPLE 1

Animals, Reagents and Antibodies

Male Sprague Dawley rats (200-250 gm) were from Sasco, Omaha, NE, and all other reagents were from Sigma Chemical company (St. Louis MO) unless otherwise stated.

Polyclonal antibodies were raised against proteins purified by immuno-affinity chromatography using reported monoclonal antibodies coupled to Sepharose 4B (Baricault et al., 1995; Sahali et al., 1988; Sahali et al., 1993). These antibodies are monospecific by immunoblotting on whole brush border preparations and by immuno precipitation of biosynthetically labeled yolk sac epithelial cells in culture (Sahali et al., 1993), and bind the cytosolic domains of the protein (Hammond et al., 1993). Control antisera included normal rabbit antiserum, and polyclonal rabbit 10 antiserum to the neurokinin-1/substance P, NK1, receptor (from Professor Jean-Yves Courard, Gif.-Sur-Yvette, France).

EXAMPLE 2

Sequencing and Estimation of N-Linked Carbohydrate of Cubilin

CNBr fragments and tryptic digests of a 100 kDa CNBr fragment of purified rat cubilin were purified by reverse phase HPLC and seven isolated peptides subjected to Edmann degradation using an Applied Biosystems 477 A sequencer equipped with a 120 A on-line chromatograph. A cross-flow reaction and the Doublot reaction and conversion cycles were used. Deglycosylation with peptide N-glycosidase F (PGNase F) of 5 μ g of purified rabbit intrinsic factor was carried out as described (Jensen et al. 1992).

25

EXAMPLE 3

cDNA Cloning, Sequencing and Northern Blotting

Total RNA was extracted from renal cortex and BN cells using Trizol (Gibco, Life Sciences) as described by the manufacturer. mRNA required for library construction was

isolated using the Qiagen Oligotex kit. Northern blots were made with 1 μ g of mRNA and revealed with 32 P labeled riboprobes (bp 1205 to 1645 and bp 1702 to 2175).

Four libraries were used. Two conventional libraries 5 were constructed in the laboratory using cDNAs synthesized by oligo dT and random priming of polyA-selected RNA from yolk sac derived BN/MSV epithelial cells using the superscript Kit (Gibco Life Sciences). After ligation to *EcoR*1 adaptors and size fractionation, they were introduced in λ Zap or λ gt11 *EcoR*1 site. 10 Subsequently screening was performed on a commercial λ Zap cDNA library (Stratagene) prepared from yolk sac derived L2 epithelial cells. Finally to identify the 5' end, a library was constructed in λ gt11 using the 5' Cap Finder library from Clontech. 15 Immunoscreening was carried out on the λ Zap-BN library using previously reported polyclonal antibodies to gp280. cDNA probes were constructed from known sequences by PCR using a 1/19 mixture of digoxigenin labeled nucleotide (Boehringer) and used to identify overlapping clones.

RACE was carried out using Marathon ready cDNA 20 prepared from rat renal cortex (Clontech). Specific primers were from bp 838-859 (SEQ ID NO. 3) for 5' RACE and bp 6872-6891 (SEQ ID NO. 4) and bp 7152-7172 (SEQ ID NO. 5) for 3' RACE. Inserts were prepared by the *ex vivo* excision system for λ zap clones (Stratagene). cDNAs from λ gt11 clones were isolated by 25 *EcoR*1 digestion and inserted in Bluescript. Sequencing was carried out by cycle sequencing in both directions with IRD-41 labeled primers and the sequence reaction were analyzed on a LICOR 4000 automatic sequencer.

EXAMPLE 4

Release of Cubilin from Renal Cortex Membranes

Rat renal cortex (0.6 g) was suspended in 3 ml PBS, pH 5 7.4, containing 0.1 mM phenylmethylsulfonylfluoride and Pefablock (Boehringer) and homogenized on ice using an ultrathorax homogenizer (23,000 rpm/min) for 20 sec. The homogenate was centrifuged at 20,800 x g for 20 min. The saline soluble and saline insoluble samples were analyzed by 10 immunoblotting with anti-cubilin and anti-megalin monoclonal antibodies (Birn et al. 1997). The amounts loaded on the gels were adjusted so that both fractions were derived from 20 µg of original cortex. IF-B₁₂ affinity chromatography of the fluid phase was performed as described (Birn et al. 1997) except that the 15 buffer contained no detergent.

Rabbit renal membranes were prepared as described (Moestrup et al. 1993). For release of cubilin, 2 mg of membranes were incubated in 525 µl of PBS, 250 units/ml heparin (LEO, Denmark), 20 mM EDTA or 5 mM phosphatidylethanolamine 20 (Sigma) for 1 h at 22°C followed by centrifugation at 20,800 x g for 20 min.

EXAMPLE 5

Immunocytochemistry

25 Rat kidneys were fixed by retrograde perfusion through the abdominal aorta with 8% paraformaldehyde in 0.1 M sodium carcodylate buffer, pH 7.2. The tissue was trimmed into small blocks, further fixed by immersion for 1 hour in the same fixative, infiltrated with 2.3 M sucrose containing 2%

paraformaldehyde for 30 minutes and frozen in liquid nitrogen. Rat embryos at day 12 of gestation were dissected free of the decidua and parietal layer to expose yolk sac epithelial cells. The tissue was then fixed by immersion and further processed as 5 described above. For electron microscopy, 70 to 90 nm cryosections were obtained at -100°C with an FCS Reichert Ultracut S cryoultramicrotome as described (Christensen et al. 1995). For double immunolabeling, the sections were incubated with the two primary antibodies overnight at 4°C after 10 preincubation in PBS containing 0.05 M glycine and 1% bovine serum albumin. Sheep anti-rat megalin serum (Moestrup et al. 1993) was diluted 1:200,000 and mouse monoclonal MAB75 (2 µg/ml) against cubilin (Sahali et al 1988). The sections were then incubated for 30 minutes with rabbit anti-sheep serum 1:20,000 15 (Dako A/S, Glostrup, Denmark), and finally incubated with 10 nm goat anti-rabbit gold particles and 5 nm goat anti-mouse gold particles (BioCell, Cardiff, UK). The sections were embedded in methylcellulose and studied in a Philips CM100 electron microscope. As controls, sections were incubated with secondary 20 antibodies alone or with non-specific monoclonal antibodies or sheep antiserum.

EXAMPLE 6

Binding of ^{125}I -Cubilin to Megalin

25 Megalin was immobilized to CNBr-Sepharose 4B (Pharmacia, Uppsala, Sweden) at a density of 0.5 mg megalin/ml gel. Cubilin was iodinated (10^6 Bq/µg) by the iodogen method (Pierce). The ^{125}I -labeled cubilin was purified by S-300 (Pharmacia) gel filtration and 10^6 cpm was loaded on the megalin-

column. After wash with binding buffer (20 mM Hepes, 150 mM NaCl, 2 mM CaCl₂, pH 7.8), bound radioactivity was eluted with the same buffer supplemented with 10 mM EDTA, counted and analyzed by SDS-PAGE.

5

EXAMPLE 7

Megalin-Cubilin Interaction Analysis by Surface Plasmon Resonance

Surface plasmon resonance measurements were 10 performed on a BIACore 2000 instrument (Pharmacia, Sweden). BIACore sensor chips (type CM5, Pharmacia) were activated with 1:1 mixture of 0.2 M N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide and 0.05 M N-hydroxysuccimine in water. Rabbit megalin was immobilized as described (Moestrup et al., 1996) at a 15 concentration of 40 µg/ml in 10 mM sodium acetate, pH 4.5 and the remaining binding sites were blocked with 1 M ethanolamine, pH 8.5. The flow buffer was 10 mM Hepes, 150 mM NaCl and 1.5 M CaCl₂, 1 mM EDTA, pH 7.4. The binding data were analyzed using the BIAscan evaluation program.

20

EXAMPLE 8

Preparation of Light Chains

Four species of light chains, two κ and two λ , were isolated and purified from the urine of four different patients with 25 myeloma, as described (Batuman et al., 1990; Batuman et al., 1997). The purity and the immunologic identity of light chains were confirmed by SDS-PAGE and Western blotting. One of the λ -light chains and the κ -light chain used here were the same light

chains used to demonstrate receptor-mediated endocytosis by radioisotope techniques. Competition experiments were initially conducted using radioiodinated λ -light chain, iodinated by the Iodobead method as reported (Batuman et al., 1990; Batuman et al., 1997). It was later switched to competition experiments with fluorescein isothiocyanate (FITC) conjugated κ -light chain. FITC conjugation was performed using FluoroTag FITC Conjugation Kit (Sigma ImmunoChemicals, St. Louis, MO).

10

EXAMPLE 9

Preparation of Renal Brush-Border Membrane Vesicles and Cortical Intermicrovillar Clefts

Rat renal cortical brush border membrane vesicles, inside/in, were isolated by magnesium precipitation technique as described (Batuman et al., 1990; Hammond et al., 1985). Rat renal cortical intermicrovillar clefts were prepared from kidneys harvested from anesthetized rats, utilizing differential Percoll gradient centrifugation and magnesium precipitation. It was also shown that the intermicrovillar clefts form vesicles "oriented inside out" *in vitro* during homogenization and can capture internally components added to the homogenization buffer.

EXAMPLE 10

Preparation of Cubilin

25 Intermicrovillar clefts prepared from renal cortices were biotinylated on the cytosolic facade using NHS-biotin (Winearls, 1995). Cubilin and the associated proteins were

purified by immunoaffinity chromatography MAb 75 was coupled to CNBr-activated Sepharose 4B (Pharmacia, Saint Quentin en Yvelines, France) as previously described (Sahali et al., 1988; 1993). Protease inhibitors were added at all steps.

5

EXAMPLE 11

Competition Between Light chains and Anti-Cubilin and Megalin Anti-Sera for Rat Renal Brush Border Membrane Binding

Binding of either [¹²⁵I]-labeled or FITC conjugated light 10 chain was investigated in the presence of up to 100,000 fold serial dilutions of anti-cubilin antibodies (Baricault et al., 1995). Equal dilutions of bovine serum albumin served as controls. With the radio-labeled light chain, binding was assayed in a gamma counter as described (Batuman et al., 1990; 1997). Binding of 15 FITC-conjugated light chain was assayed by flow cytometry using small particle techniques on a Becton-Dickinson FACStar flow cytometry with a Consort 30 computer and WinMidi software (Hammond et al., 1994; Sahali et al., 1988; 1993). The analog-to-digital conversion of fluorescence measurements on each particle 20 passes through a logarithmic amplifier such that fluorescence is expressed on a log scale.

EXAMPLE 12

Surface Plasmon Resonance Analysis of Light Chain/Cubilin Interaction

κ - or λ -light chains were immobilized *via* free amine groups to the dextran matrix of CM5 sensor chips activated by a

1/1 mixture of NHS and EDC. Unreacted sites were blocked with 1M ethanolamine, pH 8.5, (Jonsson et al., 1991; Sanders et al., 1988). The immobilization was conducted at 25°C using 10 mM HEPES, 2 mM CaCl₂, 150 mM NaCl, 0.005% NP-40, pH 7.4 as the 5 flow buffer. Then 10 mM acetate pH 4.8 was used for electrostatic pre-concentration of the protein. Different densities of κ - or λ -light chains were immobilized to three of the four flow cells; the remaining flow cell was activated and blocked with no light chains immobilized for use as a control surface. Binding experiments 10 were carried out using a BIACORE 2000 instrument.

EXAMPLE 13

Identification by Two-dimensional Gel Electrophoresis and
Microsequencing of Proteins Associated with Immunopurified
15 Cubilin

The approach used involved three steps: 1) biotinylation of intermicrovillar membranes, 2) immunoisolation of cubilin, and 3) identification of bound proteins by microsequencing. Two-dimensional electrophoresis was 20 performed according to the method of O'Farrell (O'Farrell, 1975) by Kendrick Labs, Inc. (Madison, WI). Proteins other than cubilin observed on two-dimensional gels prepared from the eluate of detergent solubilization of intermicrovillar clefts were identified by microsequencing. For this purpose, three gels were run in 25 parallel, and stained with Coomassie D. The two most abundant spots at MW 56 and 24 kDa from each gel were cut out and the material pooled. The peptides derived from the eluted proteins

by C-leu digestion were separated by HPLC, and internal peptides sequenced (Ferrara et al., 1993).

EXAMPLE 14

5 Effect of Light Chains on Endosomal Fusion

To determine if light chains had a direct effect on membrane fusion, rat renal cortical intermicrovillar clefts were prepared as described (Hammond et al., 1997; 1994) and loaded with 400 mM light chains by addition of the light chains to the 10 homogenization buffer. Fusion of these light chain loaded membranes was compared to control membranes loaded with the same concentration of albumin. All fluorescence measurements were corrected per mg of protein, and fusion reconstituted *in vitro* in cuvettes (Hammond et al., 1994). Data are expressed as mean \pm 15 standard error of the mean throughout the application. Statistical analysis was performed by analysis of variance and Bonferroni or Scheffe's post hoc comparison.

20

EXAMPLE 15

Culture of Rat Visceral Yolk Sac Cells and Internalization Experiments

The yolk sac epithelial cell line (BN/MSV) was derived from yolk sac teratocarcinoma induced by fetectomy and placental 25 injection of mouse sarcoma virus (Sahali et al., 1988). When grown under conventional conditions in modified Eagle's medium, supplemented with 2.5 mM L-glutamine, 10% fetal calf serum, and an antibiotic cocktail (penicillin, streptomycin, and Fungizone), the

cells form a domed monolayer and express abundant cubilin (Sahali et al., 1988).

EXAMPLE 16

5 Effect of Anti-Cubilin Antiserum on Endocytosis of Light Chains

Internalization experiments were conducted by exposing confluent yolk sac cells in 24-cell plates to 50 μ m FITC-conjugated light chain. These cells were selected for endocytosis experiments because cubilin expression is 100-fold greater than cultured proximal tubule cells (Sahali et al., 1988). Cells were allowed to endocytose FITC-light chain at various intervals for up to 40 minutes at 37°C with and without polyclonal anticubilin antibody at 1:1,000 dilution (added at *time 0*). This concentration is selected because it is 10-fold higher than the half-maximal inhibitory concentration of the antibody determined from the brush border binding inhibition experiments. Endocytosis is stopped by washing twice with PBS and removing light chain from medium. Cells are then trypsinized, fixed in 1% formaldehyde, and suspended in PBS, and FITC incorporated into each cell is read in a Becton-Dickinson flow cytometer as described previously. Endocytosis curves are generated by plotting fluorescence units corrected for background against time. Excess unlabeled light chain was used to test for specificity, and bovine serum albumin was used as nonspecific protein control.

25

EXAMPLE 17

cDNA Cloning of Cubilin

By immunoscreening of the λ -Zap cDNA library from rat yolk sac BN cells (Le Panse et al. 1995), an initial 0.7 kb clone

encoding a portion of cubilin was identified. The 5' sequence of this clone was used to design two nested primers to perform 5' RACE on kidney cDNA allowing identification of the 5' end of cubilin. Using PCR-generated probes for further screening of yolk 5 sac libraries a number of clones were identified. Figure 1 schematizes three overlapping clones completely sequenced and used to construct the final cDNA. The last clone contained a polyadenylation signal and a poly A tail. The 3' and 5' ends of the 11.8 kb sequence were further confirmed respectively by 10 sequencing a 3' end RACE product and a λgt11 clone selected from a Cap Finder library. Northern blot analysis of yolk sac mRNA (Figure 2) identified a mRNA of the size as the cDNA.

EXAMPLE 18

15 Primary Structure of Cubilin

The assembled cDNA (SEQ ID no. 1) revealed an uninterrupted open reading frame of 10.8 kb encoding a 20 amino acid signal peptide (in italic type) and a 3603 amino acid protein (SEQ ID NO. 2) with 42 potential N-glycosylation sites (Figure 3). 20 The molecular size of the protein backbone was calculated to 397 kDa. The seven amino acid sequences determined by N-terminal microsequencing of tryptic and CNBr peptides were all identified in the translated sequence (bolded letters in Figure 3). The size of the protein was confirmed by SDS-PAGE (Figure 4). 25 Deglycosylation of the receptor by PNGase F increased its electrophoretic mobility corresponding to a size of 400 kDa. Compared to the 460 kDa size of the untreated protein this indicates a carbohydrate content of ~13% of the receptor mass. Figure 5A shows the predicted domain organization of the

receptor. A stretch of approximately 110 amino acids with no apparent homology to known proteins is followed by a cluster of 8 EGF type B repeats which precedes 27 contiguous CUB domains accounting for 88% of the protein mass. The high degree of 5 internal homology (overall similarity of 45%) between the CUB domains is evident from the dot plot display shown in Figure 5B. A total of 76 disulfide bridges is predicted if all the extracellular modules fold normally. The only cysteine outside the CUB domains and EGF repeats, is located in the 110 amino acid N- 10 terminal sequence. This cysteine might account for the partial, disulfide bond dependent dimerization of a minor part of purified receptor (Le Panse et al., 1995; Birn et al. 1997).

Figure 6 shows alignment of the EGF repeats and CUB domains of some of the most homologous regions of other proteins. 15 Two of the EGF repeats (Nos. 2 and 4) contain the consensus sequence for Ca^{2+} binding and β -hydroxylation of Asp/Asn (Selander-Sunnerhagen et al. 1992). The 110 amino acids CUB domains contain 4 cysteines except for CUB domain 13 which is missing the first two cysteins suggested to form the upstream 20 disulfide bond (Bork and Beckmann 1993). The high homology of the CUB domains of bone morphogenic factor, the *Drosophila* dorsal-ventral patterning gene product tolloid, the embryonic protein Uvs2 in *Xenopus Laevis*, tumor necrosis factor stimulating gene 6 (Tsg6), C1r/C1s and spermidhesin is seen in the two lower 25 panels (Figure 6).

Save for the leader peptide, no sequence compatible with a transmembrane domain could be identified. This excludes the protein as a type 1 membrane protein or a glycosyl-phosphatidylinositol-anchored protein, which is synthesized with

a cleavable hydrophobic C-terminal. Furthermore, since almost the entire protein sequence consists of extracellular modules it is very unlikely that the protein is a type II or III protein with a non-cleaved hydrophobic signal peptide inserted in the membrane
5 (Levy 1996).

EXAMPLE 19

Cubilin is a Peripheral Membrane Protein

In order to verify that cubilin is a peripheral membrane protein, as predicted by the lack of a transmembrane segment and cytoplasmic tail, its release from renal cortex membranes by procedures which do not involve solubilization of the membranes or enzymatic treatment was investigated. Figure 7 shows the identical size of the renal receptor and the receptor in yolk sac and intestinal mucosa (lanes 1-3). As seen in lane 4 *vs.* lane 3, approximately 50% of cubilin was released into the fluid phase by mechanical grinding of renal cortex in PBS, whereas megalin, the 600 kDa transmembrane protein expressed in the same tissues (Saito et al. 1994), was released in minimal amounts.
10
15 Cubilin, which remained membrane-associated, was tightly bound but could be released partly by EDTA, heparin and, to a low extent phosphorylethanolamine (Figure 7, lanes 5-11). Heparin and phosphorylethanolamine have been reported to bind to the spermidhesin CUB domains (Calvete et al. 1996; Dostolova et al.,
20
25 The same treatments released virtually no megalin (Figure 7). The size of the released cubilin, as estimated by SDS-PAGE, was not different from the membrane associated cubilin.

EXAMPLE 20

Cubilin Traffics with and Binds Megalin

Previous studies have demonstrated megalin and cubilin in endocytic vesicles of the same absorptive epithelia in the intestine, kidney and yolk sac. Figure 8 shows electron microscopic examination of rat yolk sac and kidney section subjected to double gold-labeling of megalin and cubilin using a sheep anti-megalin polyclonal antibody and a mouse anti-cubilin monoclonal antibody. The large gold particles label megalin antibody and the small particles label cubilin antibody. An almost identical localization of the two sizes of gold particles was seen. Formation of cubilin/megalin complexes was tested next.

Figure 9 shows ^{125}I cubilin binds megalin covalently linked to Sepharose 4B. Bound radiolabel was released from the column by EDTA. Surface plasmon analysis (Fig. 10A) confirmed this binding. No difference in the dissociation of cubilin to megalin was seen in the pH interval 4-8. Binding of cubilin to megalin was reduced (75%) when RAP was prebound to megalin indicating that cubilin binds to the extracellular domain of megalin. Megalin-bound cubilin was still capable of binding IF-B₁₂ as shown by subjecting the megalin-chip to flow with IF-B₁₂ after the binding of cubilin (Fig. 10B). Thus, the response after adding IF-B₁₂ represents the formation of a megalin-cubilin-IF-B₁₂ complex. Control experiments showed no binding of IF-B₁₂ to megalin.

25

EXAMPLE 21

Light Chains are Ligands for Cubilin

To identify candidate ligands with which cubilin interacts, a detergent extract of rat renal apical intermicrovillar

clefts biotinylated on their cytosolic facade to affinity chromatography was subjected. The extract was passed through an immunoaffinity raised against the whole molecule. Western blot analysis of the eluate using the same antibody showed a 5 single band at the region of 460-540 kDa, consistent with cubilin (Figure 11A, left lane). Coomassie staining of a parallel gel revealed several additional bands (Figure 11A, right lane). For further characterization, the proteins eluted from the column were separated by two-dimensional gel electrophoresis and transferred, 10 and the spots were cut of the gels (Figure 11B). Pooled material representing the same spot from multiple gels was C-leu digested, fragments separated by HPLC and microsequenced (Ferrara, et al., 1993). Proteins eluted from the column included cubilin (Figure 11A, left at top of gel), a 56-kDa protein identified as the β - 15 subunit of the H⁺-ATPase by the sequence VVDLLAPYA (Figure 11B, #1), a 24-kDa protein identified as κ -light chains by the sequence (I/S)PQLLVYNA (Figure 11B, #2), and an internal tropomycin control protein added exogenously to the gel (Figure 11B, solid arrow). The 56 kDa protein was biotinylated suggesting 20 cytosolic residence, and hence was not pursued as a ligand. The 24-kDa protein was not biotinylated, suggesting exofacial residence (Table 1).

TABLE 1Analysis of Anti-gp280 Affinity Column Eluate

Protein W.M. (kDa)	Commassie Stain	Anti-gp280 Western	Cytosolic Biotinylation
540	+	+	+
56	+	-	+
24	+	-	-

5

Whether light chains are a ligand for cubilin, or were merely eluting from the antibody on the column remained uncertain. Analysis of cubilin binding to κ and λ -light chains using 10 surface plasma resonance techniques provides direct evidence that cubilin binds light chains. A stock solution of cubilin was diluted serially with flow buffer and passed over the immobilized κ -light chain surfaces for 5 minutes (50 μ l at 10 μ l/min., 25°C), followed by monitoring the dissociation phase induced by 15 introduction of cubilin free-flow buffer for 4 minutes (Figure 12A). After 4 minutes, the cubilin bound to the surface had dissociated completely, so it was not necessary to regenerate the surface prior to the next injection. The sensorgrams were corrected for bulk refractive index changes by subtracting the 20 response on the blank flowcell from the other flowcells. Cubilin bound to κ -light chains in a dose-dependent fashion (Figure 12A).

To further demonstrate the binding specificity of the cubilin to the immobilized κ -light chains, a competition

experiment was conducted. A sample of cubilin (100 nM) was incubated with κ -light chains (10 or 490 μ M), or λ -light chains (10 μ M) prior to injecting the sample over the κ -light chain surface. The binding of cubilin to the immobilized surfaces was reduced in 5 the presence of κ -light chains in a dose-response fashion (Figure 12B). Inhibition of cubilin binding to immobilized κ -light chains with 10 μ M λ -light chains suggests κ and λ light chains share a common binding site on cubilin. This series of experiments was repeated with immobilized λ -light chains, and four different light 10 chains competing (two λ and two κ) with similar results (data not shown). These studies showed that cubilin bound λ -light chains in a dose-dependent fashion, and that binding was interfered with in a dose-response fashion by both free λ - and κ -light chains. In these studies, bovine serum albumin neither competed with light 15 chains, nor bound to cubilin.

Binding of cubilin to κ -light chains was much greater at 37°C than 25°C (Figure 12C), consistent with known thermal behavior of receptor-ligand interactions (Batuman et al., 1990). Hence, BIACORE surface plasmon resonance analysis allows for 20 direct real time assay of the binding of myeloma light chains to cubilin, providing direct evidence that cubilin is a renal light chain receptor.

To determine whether light chains bind to cubilin present in brush-border membranes in its native membrane-bound form, antibody interference with light chain binding to rat 25 kidney brush-border membrane vesicles, which are known to express cubilin (Sahali et al., 1988), was tested. Binding of [¹²⁵I]-

labeled human λ -light chain to rat renal brush-border membrane vehicles is displaced by polyclonal antibodies to cubilin. The half-maximal inhibitory concentration of anti-cubilin antibody was observed at approximately 10,000 dilution (Figure 13A, solid 5 circles). In contrast, antiserum to megalin, which is known to bind these membranes (Moestrup et al., 1995), had no effect on the binding of this light chain (Figure 13A, open circles), suggesting that this λ -light chain binds exclusively to cubilin.

At the maximal inhibitory concentration, the anti-10 cubilin antiserum displacement of λ -light chain approached 90%, confirming near exclusive binding of this light chain to cubilin. It was also observed that binding of human FITC-conjugated κ -light chain to rat renal brush-border membrane vesicles was displaced by polyclonal antibodies to cubilin as assayed by flow cytometry 15 (Figure 13B). Light chain binding (45.5 ± 4.3 arbitrary fluorescent units, n=8) increased compared to unstained membranes (5.1 ± 1.2 units, n=8, p < 0.05), and was displaced by anti-cubilin (30.2 ± 1.0 units, n=8, p < 0.05). There was no effect on light chain binding by normal rabbit serum (42.9 ± 1.7 units, n=8), or antiserum to the 20 neurokinin-1/substance-P receptor (40.0 ± 1.2 units, n=4), an irrelevant antibody which binds these membranes. This provides additional evidence that the competitive effect of cubilin antiserum on the binding of light chain is specific. Flow cytometry histograms of light chain binding on a vesicle-by-vesicle basis 25 illustrate the effects of cubilin antisera on rat renal brush border binding of FITC- κ -light chains. Each histogram (Figure 13B) displays 2000 vesicles as individual dots, with FITC fluorescence plotted against vesicle size. FITC-light chains bind most but not all

brush borders (Figure 13B, left panel). Cubilin antiserum displaced FITC light chain binding (Figure 13B, right panel).

To examine the role of cubilin in light chain endocytosis, yolk sac cells were allowed to endocytose FITC-light chain in the absence and presence of anti-cubilin antiserum. These endocytosis experiments revealed a significant inhibitory effect but not total elimination of endocytosis (Figure 14). Excess unlabeled light chain and anti-cubilin antibody reduced FITC-light chain endocytosis significantly ($n=4$, $p<0.002$, Mann-Whitney-U test), whereas albumin had no effect (Figure 14A). Furthermore, a time course study showed that anti-cubilin antiserum inhibited light chain endocytosis significantly at all time intervals studied (Figure 14B, $n = 3$ each time period, $p< 0.0001$). This time course experiment also showed that anti-cubilin antiserum eliminated the saturable pattern of endocytosis with apparent linearization of the uptake curve (Figure 14B). This observation further supports that cubilin mediates light chain endocytosis in yolk sac cells. Less than complete inhibition of light chain endocytosis in the presence of anti-cubilin antiserum also indicates that, when this pathway is blocked, some light chain endocytosis occurs through alternate pathways, and that the cubilin-facilitated path is not exclusive endocytic pathway for light chains.

25

EXAMPLE 22

Function of Light Chains on Endosomal Fusion

To test whether myeloma light chains are functionally important in membrane trafficking and fusion events, intermicrovillar clefts were loaded with light chains by adding it

to the homogenization buffer (Hammond et al., 1994). Fusion reconstituted *in vitro* in cuvettes was assayed by energy transfer, and results were normalized per milligram protein (Hammond et al., 1994; Jo et al., 1995). Fusion was significantly inhibited in 5 membranes treated directly with light chains (111 ± 89 arbitrary fluorescence units/mg protein, n=8) compared with albumin entrapped controls (1584 ± 314 , n=8, $p < 0.0003$ by unpaired *t*-test, Figure 15).

10

EXAMPLE 23

Myeloma Light Chains Bind Megalin

Given the abundance of megalin on the renal brush border membrane, if megalin is a light chain receptor this would 15 predict that small polybasic drugs, such as gentamicin, which are known ligands for megalin, should compete for light chain binding to renal brush border membrane vesicles. To test this, rat renal brush border membrane vesicles were incubated in fluorescein-conjugated light chains with various concentrations of gentamicin. 20 After washing, fluorescein-light chain binding to the membranes was analyzed by flow cytometry (Figure 16). Estimate of half maximal binding concentration of gentamicin between 60 and 70 μM was placed in the middle of the curve. Gentamicin competes with fluorescent light chain binding to renal brush border 25 membranes in a dose-dependent manner (control FITC-light chain binding 39.6 ± 4.2 arbitrary fluorescence units, 10 μM gentamicin 45.1 ± 3.1 , 100 μM gentamicin $8.9 \pm 33.8^*$, 1000 μM gentamicin $6.3 \pm 1.9^*$, n=4, * $p < 0.05$). This demonstrates that gentamicin competes with light chains for brush border membrane binding.

EXAMPLE 24**Extra Renal Expression of Cubilin**

Cubilin was demonstrated to be expressed in extra-renal tissues as well (Figures 17A and 17B). Spleen, brain, liver, 5 heart, kidney and thyroid are the possible sites where cubilin is expressed. Administering cubilin in a pharmaceutically acceptable carrier might lead to the reduction of toxicity and therefore protecting those sites.

10

EXAMPLE 25**Urine Cubilin**

Cubilin was also detected in the urine (Figure 18), which indicates that cubilin is released into the urine and the 15 assay of urinary cubilin might be an excellent marker for detecting renal damage. Several conditions can be considered: 1) renal damage of acute origin may increase the excretion of a tubular protein such as cubilin and constitute a more sensitive and specific marker than adenosine deaminase (Iglesias et al., 1994; 20 Parvez et al., 1990; Tolkoff-Rubin et al., 1987); 2) chronic renal damage with tubular atrophy may be associated on the other hand with a reduced expression of cubilin, the assessment/follow up of which may be useful to monitor the evolution of renal fibrosis; 3) because Imerslund Grasbeck (IG) patients do not all have 25 proteinuria, it is likely that the syndrome is associated with various mutations of cubilin (for comparison one or two hundred have been described in familial hypercholesterolemia) which may be associated with variable levels of cubilin excretion. In fact the assay of IF-B₁₂ binding activity of the urine of IG patients has

been studied (Dugue et al., 1998; Gueant et al., 1995); and 4) some diseases characterized by proteinuria of unknown cause may be due to a defect in cubilin, which may be absent and therefore not detectable in the urine, or excreted in large amounts if it is not gathered adequately to the membrane.

More generally it can be presumed that mutations of key receptors of trophoblastic cells, such as cubilin may account for a variety of pathologies. For instance, the bulk of fetal malformations, not accounted for by the known or suspected 10 hereditary abnormalities, might be related to cubilin defect. Similarly, most cases of poor fetal development or fetal loss, which are of unknown origin up to date, might be caused by cubilin mutations.

The present study provides novel molecular 15 information on cubilin, previously known as the yolk sac target antigen of teratogenic antibodies and the intestinal receptor for IF-B₁₂. The primary structure predicts 35 extracellular modules uniquely organized in a cluster of 8 EGF repeats followed by, from a molecular point of view, an huge cluster of 27 CUB domains 20 which account for 88% of the mass of the protein. Northern and western blotting of kidney, yolk sac and intestine indicate no difference in size of the receptor in these organs.

The EGF type B repeats are similar to the carboxyl-terminal extracellular modules of megalin and low density 25 lipoprotein receptor-related protein. Cubilin has otherwise very little homology to these two giant receptors, which also bind RAP and mediate endocytosis of a variety of ligands. Also cubilin does not display homology to sortilin, the 95 kDa putative vesicular sorting receptor, which also binds RAP (Petersen et al., 1997). The

CUB domains conform to the description of Bork and Beckmann (1993) based on the analysis of 31 copies of a module initially identified in C1r and C1s components of complement and subsequently in a variety of proteins associated with fetal 5 development. They consist of 110 amino acids defining a characteristic hydrophobicity pattern predicted to form antiparallel beta barrels (Dias et al. 1997). The four conserved cysteines, generally thought to form two S-S bridges (1-2, 3-4), are found in all but domain 13 of cubilin which lacks the first 2 10 cysteines as already described in the first CUB domains of C1r/s and the homologues MASP1/2. When analyzed individually, the CUB domains of cubilin are more closely related to those seen in developmental control proteins.

On the functional level there is compelling evidence 15 that the CUB domains are involved in the binding of proteins, as described for the Ca^{2+} -dependent formation of the C1 complex (Bosby and Ingham 1990), as well as for binding of phospholipids and carbohydrates, as demonstrated for spermidhesins (Calvete et al. 1996a and 1996b, Dostalova et al. 1995). In addition to the 20 CUB domains, the EGF repeats might also account for some of the binding properties of cubilin. EGF repeats are widely expressed and involved in a number of receptor-ligand interactions (Davis 1990). Two of the EGF repeats in cubilin have the consensus sequence for calcium binding (Rao et al. 1995) and may be 25 involved in the calcium-dependent binding of e.g. RAP or IF-B₁₂.

The lack of a transmembrane segment was surprising since in previous studies isolation of cubilin relied on the use of detergent solubilized membranes. Furthermore the previous studies showed that cubilin was internalized through clathrin

coated organelles and recycled *via* dense apical tubules (Le Panse et al. 1995). This first suggested to reassess membrane tethering of cubilin. Early results was thus confirmed indicating that an intrinsic factor-B₁₂ binding protein (Cotter and Rothenberg 1976) 5 and the target of teratogenic antibodies (Leung 1982) could be released at least in part from intestinal or renal tissue using mechanical dissociation in the absence of detergents. It was further showed that whereas membrane association was stable between pH 4 and 8 cubilin could be released by heparin, 10 phosphorylethanolamine and EDTA. These observations, which indicate nonionic interactions with sugars and phospholipids, are in line with the membrane binding properties of spermadhesins which consist of a single CUB domain, lack a transmembrane segment, but are yet tightly bound to the surface of sperm cells 15 *via* phospholipids (Dostalova et al. 1995). Another region of the same CUB domain binds to carbohydrates of zona pellucida, the extracellular investment surrounding the mammalian egg. The lectin binding characteristics of the spermadhesins are not fully characterized but include heparin and a variety of carbohydrates 20 including Gal beta (1-4)-GlcNac and Gal beta (1-3)-GlcNac (Calvete et al. 1996,1997). In view of the 27 CUB domains present in cubilin this receptor may have multiple membrane attachments which may account for the inability to release all the membrane associated cubilin.

25 The identification of the components which link cubilin to the membrane is also essential for understanding its internalization and recycling. The present study suggested that the binding of cubilin to megalin is crucial for this process. Co-internalization of a receptor which lacks internalization signal(s)

by means of another receptor has previously been shown. The GPI linked urokinase receptor can thus be endocytosed by coupling of urokinase receptor-bound urokinase/inhibitor complex to the LDL receptor-related protein (Nykjær et al. 1997, Conese et al. 1996). It is likely that a similar process can be mediated by megalin which can also bind the urokinase-inhibitor complex (Moestrup et al. 1993). Based on the strict colocalization of cubilin and megalin at the subcellular level and on the ability of megalin to bind cubilin *in vitro*, it was proposed that megalin is crucial for the internalization of cubilin and cubilin ligand complexes. After internalization, the ligand IF-B₁₂ is segregated from the receptor and directed to lysosomes for degradation of IF (Dan and Cutler 1994, Birn et al. 1997) whereas cubilin is recycled to the membrane. Since the cubilin/megalin complex is stable at pH 5, the two receptors might remain in complex during the entire recycling pathway at variance from the urokinase receptor which recycles to the plasma membrane without being linked to the LDL receptor-related protein (Nykjær et al. 1997).

Upon analysis of the effect of polyclonal megalin antibodies and RAP on the endocytosis of ¹²⁵I-IF-B₁₂ in cultured yolk sac cell, only a 10 to 15% reduction was found in uptake. This modest effect might be accounted for a short cell surface expression of megalin and cubilin due to rapid recycling of the two proteins and thereby a too short time for the cubilin-megalin to dissociate, a prerequisite for RAP to block binding. Furthermore, a continuous incubation with RAP will probably have no effect on intracellular receptors, since externally receptor-bound RAP is transported to lysosomes for degradation (Iadonato et al., 1993).

In order to further characterize the partnership of these two giant receptors, studies have been initiated to investigate cubilin trafficking in megalin deficient or megalin-mutated cells expressing cubilin. However, such analysis might be 5 complex since recent data on megalin deficient mice indicate a key role of megalin for normal development of the endocytic apparatus in the proximal kidney tubules and for survival of the mice in general (Willnow et al. 1996).

The observation that the target of teratogenic 10 antibodies contains CUB domains is of particular interest in view of the fact that these domains are often observed in developmentally regulated proteins. The mode of action of the teratogenic antibodies is not known but have been shown to inhibit endocytosis, thus reducing the amount of maternal proteins 15 internalized and consequently the amount of protein derived amino acid which can be incorporated into embryonic tissue (Beckman et al., 1997; Lloyd, 1990, Le Panse et al., 1994). However, there is no direct evidence that a decreased amino acid 20 supply is responsible for foetal malformations. Alternatively, the teratogenic effect might relate to a more specific disturbance of the materno-foetal barrier such as an impaired transfer of B₁₂ or 25 of other nutrients. Interestingly, the pattern of antibody induced fetal malformations which includes abnormal cranio-facial development, in particular of the eyes and hypophysis (Sahali et al., 1988) resembles to some extent the holoprosencephalic syndrome induced by anti-cholesterol agents (Llirbat et al. 1997), knock out of the cholesterol-depedent Sonic hedgehog (shh) gene (Porter et al., 1996) or of the megalin gene (Willnow et al., 1996). It has been proposed (Herz et al., 1997) that the defective

development of the central nervous system in megalin-deficient mice was related to a decreased megalin-mediated uptake of cholesterol-containing lipoproteins which in turn altered the addition of cholesterol to the shh protein. It is therefore possible 5 that anti-cubilin antibodies could interfere with cholesterol uptake either directly or indirectly *via* binding of cubilin to megalin in the yolk sac.

The present study establishes cubilin as a novel type of peripheral membrane receptor with multiple potential sites for 10 interaction with other proteins and membrane components. Cubilin can bind IF-B₁₂, RAP, megalin, and most likely calcium, phospholipids and carbohydrates (Table 2). However a number of ligands may remain to be identified in order to explain the role of the receptor in kidney function and its importance in fetal 15 development.

TABLE 2Ligands for Cubilin

Ligand	BN cell uptake	surface plasmon resonance	antibody interference	competition with known ligands
light chains*	X	X	X	X
myoglobin		X	X	X
metallothionein		X	X	X
haptoglobin			X	X
polybasic drugs		X	X	X
Intrinsic factor vitamin B ₁₂		X	X	X
LDL	X			X
HDL	X			X
transferrin	X			
RAP		X	X	X
albumin	X	X	X	X

* as light chains compete with renal brush border membrane binding with multiple other medically and physiologically important proteins,

5 this suggests that these proteins are also ligands for cubilin. This group of proteins include β 2 microglobulin, amyloid, insulin, cytochrome c and interferon. LDL: Low density lipoprotein; HDL: High density lipoprotein; RAP: Receptor related protein.

Solution structure of spermadhesin PSP-I/PSP (Romero et al. Nat. Struct. Biol 1997 10:78-788), a dimer which consists exclusively of two single CUB domains has recently been obtained. 5 It reveals that CUB domains are characterized by 2 layers of 5 beta sheets, the top layer of one of the CUB domains contacting the lower layer of the other CUB domain in a manner that leaves preferentially exposed the less conserved beta turns which carry the ligand binding sites. If such an arrangement prevails in 10 cubilin, it is likely to account for and one would predict a wide variety of ligands binding to distinct CUB domains.

Indeed the fact that RAP binds cubilin but does not inhibit binding of IF-B₁₂ complexes indicates that these 2 ligands bind to different sites and probably to different CUB domains. 15 Similarly haptoglobin and light chains both bind cubilin but do not compete for binding. The modular structure of cubilin thus strongly suggests that it may be feasible to produce fragments of small size corresponding to one or a few CUB domains which can be used therapeutically: this type of fragments will bind selected 20 ligands but preserve many/most other functions of cubilin.

The present studies also showed that cubilin, a giant receptor which participates in the endocytic scavenger pathway of the renal proximal tubule cells, binds and facilitates endocytosis of immunoglobulin light chains isolated from the urine of myeloma 25 patients. Evidence that cubilin is a light chain receptor came from the analysis of eluates from an affinity column prepared with anti-cubilin antiserum in which cubilin coeluted with κ -light chain. The κ -light chain was definitively identified by microsequencing

after isolation by two-dimensional electrophoresis. Several additional lines of evidence add weight to the hypothesis that cubilin is a light chain receptor. Competition experiments by anti-cubilin antiserum and surface plasmon resonance experiments 5 both showed that all tested light chains bind to cubilin.

Surface plasmon resonance technology allowed direct analysis of the binding of light chains to cubilin. Several characteristics of the observed sensorgrams suggest that light chains bind cubilin specifically. First, cubilin bound to light chains 10 in a temperature- and dose-dependent manner whether κ or λ -light chain is immobilized. Second, four species of non-immobilized light chains all interfered with binding in a dose-dependent manner. Third, the kinetics of binding and displacement were very similar to values reported using 15 radioactive membrane binding techniques (Batuman et al., 1997; Driesbach et al, 1994; Marchalonis et al., 1992). Lastly, λ -light chains interfere with κ -light chain binding to cubilin and vice versa. This data revalidates the use of surface plasmon resonance technology to quantitate low affinity binding (Jonsson et al., 1991; 20 Sanders et al., 1988).

As κ -light chains are 100-fold more abundant than λ -light chains in healthy animals and humans (Riedel et al., 1991), it is not surprising to observe κ -light chains eluting from the cubilin affinity column but not λ -light chains. The current surface 25 plasmon resonance data provides direct evidence confirming and extending the observation made by membrane binding of light chains: both κ - and λ -light chains are ligands for cubilin.

Studies of classic binding kinetics utilizing Scatchard analysis demonstrate several ligands competing with light chain for brush border membrane binding. These ligands include lysozyme, insulin, cytochrome c, myoglobin and β_2 -microglobulin 5 (Batuman et al., 1990; 1997; Driesbach et al., 1994). Competition by low molecular weight proteins raise the probability that cubilin is a multi-ligand receptor responsible for the endocytosis and cellular trafficking of a number of proteins normally filtered in the glomerulus and catabolized in the kidney, extending the role 10 of this scavenger pathway receptor to such diverse phenomena as rhabdomyolysis and insulin metabolism. The multiple putative ligands for cubilin reflect the precedent set by other giant glycoprotein receptors such as the low-density-lipoprotein receptor, megalin, and the α_2 -macroglobulin receptor, which bind 15 many ligands with a spectrum of affinities at multiple binding sites (Moestrup et al., 1994). Cloning data that reveal multiple EGF repeats and CUB domains strengthens this expectation.

Receptor kinetic studies have demonstrated that light chain binding to receptors in cultured proximal tubule cells is 20 followed by endocytosis and ultimate lysosomal degradation. The present observations suggest that cubilin is a receptor that can mediate endocytosis of light chains in renal proximal tubular cells. Nearly 90% of the λ -light chain binding was displaced by anti-cubilin antibody. In contrast, anti-megalin antibody did not 25 compete with the brush border binding of this light chain at all. This suggests that cubilin is the quantitatively major receptor for this λ -light chain. However, at maximal inhibitory concentration of the anti-cubilin antibody, 10% of light chain remained bound to

brush border membranes, suggesting presence of additional binding sites for this light chain.

Anti-cubilin antiserum also inhibited endocytosis of light chain significantly. This further confirms that cubilin binding is followed by endocytosis of light chain. However, less than total inhibition of light chain endocytosis by anti-cubilin antibody indicates that this pathway may not be the exclusive endocytic pathway for light chains and that there may be alternate pathways which can compensate partially when the cubilin-mediated pathway is blocked. The antibodies used in the present study may be less than blocking functionally, and incomplete inhibition of endocytosis may be on this basis.

Importantly, binding of light chains to scavenger pathway receptors is not just a structural observation, as light chains had potent direct effects on endosomal fusion reconstituted *in vitro*. This raises the novel hypothesis that ligand binding may affect fusion properties of membranes, mediated by the receptors they bind. Select ligands are known to induce endocytosis of the ligand-receptor complex by binding, and the protein components of the final common pathway of fusion have largely been identified and cloned. These mechanisms may provide new insights into nephrotoxicity of myeloma light chains and other nephrotoxic low molecular weight proteins.

There are 13,500 new cases of myeloma annually in the U.S. and 1-4 new cases/100,000 of population worldwide. Although the precipitation of light chains with Tamm-Horsfall protein to form casts in renal distal nephron segments has been defined down to specific peptide sequences, the molecular

characteristics of receptors that mediate the endocytosis of light chains in the proximal tubule have not been defined. Identification of the proximal tubular receptor for light chains extends and compliments these observations. The proximal tubule 5 determines the distal delivery of low molecular weight proteins by reabsorbing the bulk of filtered proteins including light chains. Many low-molecular weight proteins induce injury to the proximal tubule, while others precipitate in the distal nephron. Both these mechanisms contribute to the pathogenesis of tubulointerstitial 10 nephropathies associated with low-molecular-weight proteins, such as multiple myeloma. Proximal reabsorption of light chains is associated with tubular atrophy, necrosis and Fanconi syndrome. Taken together with understanding of distal tubular cast formation, identification of major renal binding proteins for 15 myeloma light chains in the proximal tubules will allow detailed characterization of the binding site between cubilin and light chains, as well as other nephrotoxic low-molecular weight proteins. This completes the necessary mechanistic data of all affected nephron sites for the rational design of agents to protect 20 from nephrotoxicity caused by myeloma light chains as well as other low-molecular weight proteins.

The present studies also demonstrate several lines of evidence suggesting that light chains are a ligand for megalin: anti-megalin antiserum partially displaces brush border light 25 chain binding, and gentamicin displaces brush border light chain binding. Independent evidence suggests that cubilin is a receptor for polybasic drugs as gentamicin directly interferes with light chain binding to cubilin *in vitro*. These observations are important, both to understand the complex interactions of toxic

and physiological ligands on proximal tubule scavenger pathway receptors, as well as the eventual development of clinical protective agents for nephrotoxic damage mediated by ligands for cubilin and/or megalin.

5 The following references were cited herein.

Baricault et al., *Biochem Biophys Res Commun* 212:353-359, 1995.

Batuman et al., *Am J Physiol* 258 (Renal Fluid Electrolyte Physiol 27): F1259-F1265, 1990.

Batuman et al., *Am J Physiol*. 272 (Renal Fluid Electrolyte Physiol 41):F521-F530, 1997.

Bennett et al., *Proc. Natl. Acad. Sci. USA* 90: 2559, 1993.

Bork et al., *J Mol Biol.* 231, 539-45, 1993.

Brent et al., *Proc. Soc. Exp Biol. Med.* 106:523-526, 1961.

Busby et al., *Biochemistry* 29, 4613-4618, 1990.

15 Bu et al., *J. Biol. Chem.* 271, 22218-24, 1996.

Calvete et al., *Biol. Chem.* 377, 521-527, 1996.

Conese et al., *J. Cell Biol.* 131, 1609-1622, 1996.

Cotter et al., *J. Haematol.* 34, 477-87, 1976.

Dan et al., *J. Biol. Chem.* 269, 18849-18855, 1994.

20 Davis, *The New Biologist* 2, 410-419, 1990.

Dias et al., *Protein Sci.* 6, 725-7, 1997.

Dostalova et al., *Biol. Chem. Hoppe Seyler* 376, 237-242, 1995.

Driesbach et al., *Renal Physiology* 17:287-293, 1994.

Dugue et al., *J Pediatr Gastroenterol Nutr* 26(1):21-5, 1998.

25 Farquhar et al., *J. Am Soc. Nephrol.* 6, 35-47, 1995.

Ferrara, et al., *Chemistry IV*" (ed. R Angeletti) p 379-387, 1993.

Fyfe et al., *J. Biol. Chem.* 266, 4489-94, 1991.

Grasbeck et al., *Acta Med. Scan.* 167, 289-296, 1960.

Gueant et al., *Gastroenterology* 108, 1622-1628, 1995.

Hishida et al., *EMBO J.* 15, 4111-22, 1996.

Hammond et al., *Kidney Int* 42:997-1005, 1992.

Hammond et al., *Am J Physiol* 272 (Renal Fluid Electrolyte Physiol 41) :F117-F123, 1997.

5 Hammond et al., *Cytometry* 14: 411-420, 1993.

Hammond et al., *Am J Physiol* 267 (Renal Fluid Electrolyte Physiol 36):F1021-F1033, 1994.

Hammond et al., *Am J Physiol* 267:F516-F527, 1994.

Hammond et al., *J Clin Invest* 75:1983-1989, 1985.

10 Huang et al., *J Clin Invest* 99(4):732-736, 1997.

Iglesias et al., *Transplant Proc.* 26(1):75-6, 1994.

Imerslund, *Acta Paediatr. Scand.* 49 (Suppl), 1-115, 1960.

Jensen et al., *FEBS Lett.* 29, 129-32, 1992.

Jo et al., *Proc Natl Acad Sci USA* 92:1876-1880, 1995

15 Jonsson et al., *Biotechniques* 11:620-627, 1991.

Krieger et al., *Ann. Rev. Biochem.* 63, 601-637, 1994.

Le Panse et al., *Am. J. Pathol.* 145, 1526-36, 1994.

Le Panse et al., *Eur. J. Cell Biol.* 67, 120-9, 1995

Levi, *Essays in Biochemistry* 31, 49-59, 1996.

20 Llirbat et al., *J. Lipid Res.* 38:1 22-34, 1997.

Leung et al., *J. Exp. Med.* 156, 372-84, 1982.

Marchalonis et al., *J Prot Chem* 11(2):129-137, 1992.

Moestrup et al., *J. Biol. Chem.* 268,16564-16570, 1993.

Moestrup et al., *Biochimica et Biophys. Acta* 1197:197-213, 1994

25 Moestrup et al., *J Clin Invest* 96:1404-1413, 1995.

Nykjær et al., *EMBO J.* 16, 2610-2620, 1997.

O'Farrel P.H, *J. Biol Chem* 250:4007-4021, 1975.

Parvez et al., *Clin Chim Acta* 190(1-2):111-3, 1990.

Rao et al., *Cell* 82, 131-141, 1995.

Riedel et al., From: "Neoplastic diseases of the blood" Second Ed.
Eds: Wiernik et al., Churchill Livingstone, 1991, Chap. 23:347-372.

Rothman et al., *Science* 272:227-234, 1996.

Rothman et al., *Current Biology* 4(3):220-233, 1994.

5 Sahali et al., *J Exp Med* 167:213-218, 1988.

Sahali et al., *Am J Pathol* 142:1654-1667, 1993.

Saito et al., *Proc Natl Acad Sci USA* 91:9725-9729, 1994.

Sanders et al., *J. Clin. Invest.* 82:2086-2096, 1988.

Tolkoff-Rubin et al., *Nephrol Dial Transplant* 2(3):143-8, 1987.

10 van der Merwe et al., *Curr Opin in Immunol* 8:257-261, 1996.

Yamamoto et al., *Cell* 39:27-38, 1984.

Willnow et al., *EMBO J.* 15, 2632-2639, 1996.

Winearls, *Kidney International* 48:1347-1360, 1995.

Zeidel et al., *Am J Physiol* 265:F822-F833, 1993.

15 Any patents or publications mentioned in this specification are indicative of the levels of those skilled in the art to which the invention pertains. These patents and publications are herein incorporated by reference to the same extent as if each individual publication was specifically incorporated by reference.

20 One skilled in the art will readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The present examples along with the methods, procedures, treatments, molecules, and specific compounds

25 described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention as defined by the scope of the claims.

WHAT IS CLAIMED IS:

1. DNA encoding a cubilin protein selected from the group consisting of:

5 (a) isolated DNA which encodes a cubilin protein;
(b) isolated DNA which hybridizes to isolated DNA of
(a) above and which encodes a cubilin protein; and
(c) isolated DNA differing from the isolated DNAs of
(a) and (b) above in codon sequence due to the degeneracy of the
10 genetic code, and which encodes a cubilin protein.

2. The DNA of claim 1, wherein said DNA has the sequence shown in SEQ ID No. 1.

15 3. The DNA of claim 1, wherein said cubilin protein has the amino acid sequence shown in SEQ ID No. 2.

20 4. The DNA of claim 1, wherein said DNA is expressed in the tissues selected from the group consisting of kidney, spleen, brain, liver, heart and thyroid.

25 5. A vector capable of expressing the DNA of claim 1 adapted for expression in a recombinant cell and regulatory elements necessary for expression of the DNA in the cell.

6. The vector of claim 5, wherein said DNA encodes a cubilin protein having the amino acid sequence shown in SEQ ID No. 2.

7. A host cell transfected with the vector of claim 5, said vector expressing a cubilin protein.

8. The host cell of claim 7, wherein said cell is
5 selected from the group consisting of bacterial cells, mammalian cells, plant cells and insect cells.

9. The host cell of claim 8, wherein said bacterial cell is *E. coli*.

10

10. Isolated and purified cubilin protein or fragment coded for by DNA selected from the group consisting of:

(a) isolated DNA which encodes a cubilin protein or fragment;

15 (b) isolated DNA which hybridizes to isolated DNA of (a) above and which encodes a cubilin protein or fragment; and

(c) isolated DNA differing from the isolated DNAs of (a) and (b) above in codon sequence due to the degeneracy of the genetic code, and which encodes a cubilin protein or fragment.

20

11. The isolated and purified cubilin protein or fragment of claim 10, wherein said cubilin protein having the amino acid sequence shown in SEQ ID No. 2, wherein said fragment having amino acid sequence consisting of one or more of the 25 sequences selected from the group consisting of SEQ ID Nos. 21-27.

12. A method of detecting expression of the protein or fragment of claim 10 in a sample, comprising the steps of:

(a) contacting mRNA obtained from said sample with a labeled hybridization probe; and

(b) detecting hybridization of said probe with said mRNA.

5

13. A pharmaceutical composition comprising the protein or fragment of claim 10 and a pharmaceutically acceptable carrier.

10 14. A method of treating or reducing toxicity in an individual in need of such treatment, comprising the step of:

contacting said individual with the pharmaceutical composition of claim 13.

15 15. The method of claim 14, wherein said toxicity occurs in the tissues selected from the group consisting of kidney, spleen, brain, liver, heart and thyroid.

20 16. A receptor for ligands, wherein said receptor comprising a cluster of EGF repeats and a cluster of CUB domains.

17. The receptor of claim 16, wherein said receptor is cubilin.

25 18. The receptor of claim 16, wherein said ligand is selected from the group consisting of immunoglobulin light chain, myoglobin, intrinsic factor-vitamin B₁₂, metallothionein, β-2-microglobulin, amyloid, hemoglobin, haptoglobin, interferon, insulin, cytochrome c, lysozyme, transferrin, transthyretin,

polybasic drugs, low density lipoprotein, high density lipoprotein and receptor related protein.

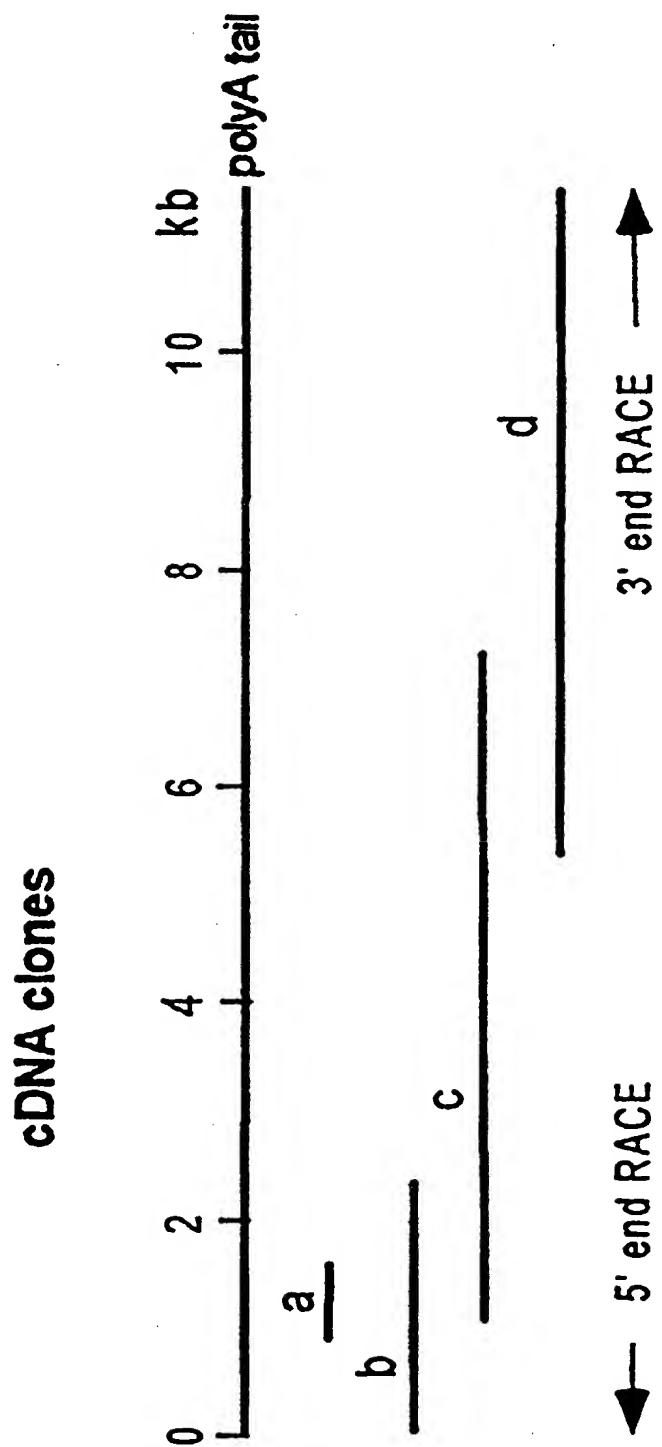
19. The receptor of claim 18, wherein said
5 immunoglobulin light chain is selected from the group consisting
of κ -light chain and λ -light chain.

20. A method of detecting renal damage in an
individual suspected to have renal damage, comprising the steps
10 of:

extracting a urine sample from said suspected
individual and a urine sample from a normal individual;

measuring level of cubilin in said urine samples;

15 comparing the level of cubilin in said suspected
individual with the level of cubilin in said normal individual,
wherein if said level of cubilin in said suspected individual is
lower than said level of cubilin in said normal individual, said
suspected individual has chronic renal damage; wherein if said
20 level of cubilin in said suspected individual is higher than said
level of cubilin in said normal individual, said suspected
individual has renal damage of acute origin.



1/23

SUBSTITUTE SHEET (RULE 26)

FIG. 1

kb

- 9

- 6

- 5

- 4

- 3

- 2

- 1.5

FIG. 2

2/23

-20 MSSQFLWGFV TLLMIAELDG KTGKPEQRGQ KRIADLHQPR MTTEEGNLVF
 31 LTSSTQNI EF RTGSLGKIKL NDEDLGECLH QIQRNKDDII DLRKNTTGLP
 81 QNILSQVHQL NSKLVDL E RD FQNLQONVER KVCSSNPCLN GGTCVNLHDS
 131 FVCICPSQWK GLFCSEDVNE CVVYSGTPFG CQSGSTCVNT VGSFRCDCTP
 181 DTYGPQCASK YNDCEQGSQKQ LCKHGICEDL QRVHHGQPNF HCICDAGWTT
 231 PPNGISCTED KDEC SLQPS CSEHAQCFNT QGSFYCGACP KGWQGNGYEC
 281 QDINECEINN GGCSQAPIVP CLNTPGSFSC GNCPAGFSGD GRVCTPVDIC
 331 SIHNGGCHPE ATCSSPVLG SFLPVCTCPP GYTGNGYGSN GCVRLSNICS
 381 RHPCVNGQCI ETVSSYFCKC DSGWSGQNCT ENINDCSSNP CLNGGT CIDG
 431 INGFTCDCTS SWTGYYCQTP QAACGGILSG TQGTFAYHSP NDTYIHNVNC
 481 FWIVRTDEEK VLHVTFTFFD LESASNCPRE YLQIHGDSS ADFPLGRYCG
 531 SRPPQGIHSS ANALYFHLYS EYIRSGRGFT ARWEAKLPEC GGILTDNYGS
 581 ITSPGYPGNY PPGRDCVWQV LVNPNSLITF TFGTLSLEH NDCSKDYLEI
 631 RDGPFHQDPV LGKFC TSLST PPLKTTGPAA RIHFHSDSET SDKGFHITYL
 681 TTQSDLDCCG NYTDTDGE LL PP LSGPF SH SRQC VYLITQ AQGEQIVIN
 731 THVELESQMG CSHTYIEVGD HDSLLRKICG NETLPIRSV SNKVWIRLRI
 781 DALVQKASFR ADYQVACGGM LRGE GFFRSP FYPNAPGRR TCRWTISQPO
 831 RQVVLNFTD FQIGSSASCD TDYIEIGPSS VLGSPGNEKF CSSNIPSFIT
 881 SVYNILYVTF VKSSSMENRG FTAKFSSDKL ECGEVLTAST GIIESPGHPN
 931 VYPRGVNCTW HVVVQRGQLI RLE FSSFYLE FHYNCTNDYL EIYDTAAQTF
 981 LGRYCGKSIP PS LTSNSNSI KLIFVSDSAL AHEGFSINYE AIDASSVCLY
 1031 DYT DNF GMLS SPNFPNNYPS NWECIYRITV GLNQ QIALHF TDFTLEDYFG
 1081 SQCVDFVEIR DGGYETSP LV GIYCGSVLPP TIISHSNKLW LKFKSDAALT
 1131 AKGFSAYWDG SSTGCGGNLT TPQVLTSPNY PMPYHHSSEC YWRLEASHGS
 1181 PFELEFQDFH LEHHPSCSLD YLGRV DGP TT NSRLIDKLCG DTT PAPIRSN
 1231 KDVVLLKTEE LMQGQLGRGF EINFRQRCDN VYIVNKTFGI LESINYPNPY
 1281 DKNQRCNWTI QATTGNTVNY TFLGF DVES Y MNCSTDYVEL YDGPQWMGRY
 1331 CGNNMPPPGA TTGSQLHVL F HTD GINS GEK GFKMQWFTHG CGGEMSGTAG
 1381 SFSSPGYPNS YPHNKECIWN IRVAPGSSI Q LTIHDFDVEY HTSCNYDSLE
 1431 IYAGLDFNSP RIAQLCSQSP SANPMQVSST GNE LAIRFKT DSTLN RGFN
 1481 ASWRAVPGC GGIIQLSRGE IHSPNYPNNY RANTECSWII OVERHHRVLL
 1531 NITDFDLEAP DSCLRLMDGS SSTNARVASV CGRQQPPNSI IASGN SLFVR
 1581 FRSGSSSQNR GFRAE FREEC GGRIMTDSSD TIFSPLYPHN YLHNQNC SWI
 1631 IEAQPPFNHI TLSFTHFQLQ NSTDCTRDFV EILDGNDYDA PVQGRYCGFS
 1681 LPHPIISFGN ALT VRFVTDS TRSFEGFRAI YSASTSSCGG SFYTL DGIFN
 1731 SPDYPADYHP NAE CVWNIA SPGNRLQLSF LSFNLENSLN CNKDFVEIRE
 1781 GNATGHLIGR YCGNSLPGNY SSAEGHSLWV RFVSDGSGTG MGFQARFKNI
 1831 FGNNNIVGTH GKIAPFWPG KYPYNSNYKW VVNVDAYHII HGRILEMDIE
 1881 PTTNCFYDSL KIYDGFDTHS RLIGTYCGTQ TESFSSSRNY LTFQFSSDSS
 1931 VSGRGFLLEW FAVDVSDSTP PTIAPGACGG FMVTGDTPVH IFSPGWPREY
 1981 ANGADCIWII YAPDSTV E LN ILSLDIEPQQ SCNYDKLIVK DGDSDLSP EL
 2031 AVLCGVSPPG PIRSTGEYMY IRFTSDTSVA GTGFNASFHK SCGGYLHADR
 2081 GVITSPKYPD TYLPNLNC SW HVLVQTGLTI AVHFEQPFQI QNRDSFCSQG
 2131 DYLVLRNGPD NHSPLGP SG RNRFCGMYA PSTLFTSGNE MFVQFISDSS

FIG. 3-1

3/23

2181 NGGQGFKIRY EAKSLACGGT VYIHDADSDG YLTSPNYPAN YPQHAECIWI
 2231 LEAPPGRSIQ LQFEDQFNIE DTPNCVSYI ELRDGANSNA RLVSKLCGHT
 2281 LPHSWVSSRE RIYLKFTDG GSSYMGFKAK YSIASCGBTV SGDSGVIESI
 2331 GYPTLPYANN VFCQWFIRGL PGHYLTLSFE DFNLQSSPGC TKDFVEIWEN
 2381 HTSGRVLGRY CGNSTPSSVD TSSNVASVKF VTDGVTASG FRLQFKSSRQ
 2431 VCGGDLHGP GTFTSPNYPN PNPHARICEW TITVQEGRRI VLTFTNLRIS
 2481 TQPSCNSEHL IVFNGIRSN S PLLQKLCSRV NVTNEFKSSG NTMKVVFFTD
 2531 GSRPYGGFT SYTSTEDAVC GGFLPSVSGG NFSSPGYNGI RDYARNLDCE
 2581 WTLSNPNREN SSISIYFLEL SIESHQDCTF DVLEFRVGDA DGPLIEKFCS
 2631 LSAPTAFLVI PYPQVWIRFV SNERVEYTGF YIEYSFTDCG GIRTGDNGVI
 2681 SSPNYPNLYS AWTCSWLLK APEGHTITLT LSDFLLEAHP TCTSDSVTVR
 2731 NGDSPGPVVI GRYCGQSVPR PIQSGSNQLI VTFNTNNQGQ TRGFYATWTT
 2781 NALGCGGTFH SANGTIKSPH WPQTFPENSR CSWTVITHDS KHWEISFDSN
 2831 FRIPIPSSDSQC QNSFVKVWGG RLMINKTLLA TSCGDVAPSP IVTSGNIFTA
 2881 VFQSEEMAAQ GFSASFISRC GRTFNTSPGD IISPNFPKQY DNNMNCTYLI
 2931 DADPQSLVIL TVFSFHLEDR SAITGTCDHD GLHIIKGRNL SSTPLVTICG
 2981 SETLRPLTVI GPVLLNFYSD AYTTDFGFKI SYRAITCGGI YNESSGILRS
 3031 PSYSY\$NYPN NLYCVYSLHV RSSRVIIRF NDFDVAPSNI CAHDFLEVFD
 3081 GPSIGNRSILG KFCGSTRPQT VKSTNSSLTL LFKTDSSQTA RGWKIFFRET
 3131 IGPQQGCGGY LTEDNQSFVS PDSDSNGRYD KGLSCIWYIV APENKLVKLT
 3181 FNVFTLEGPS SAGSCVYDYY QIADGASINS YLGGKFCGSR MPAPFISSGY
 3231 FLTFQFVSDV TVEMRGFNAT YTFVDMPCGG TYNATSTPQN ASSPHLSNIG
 3281 RPYSTCTWVI AAPPQQQVQI TVWDLQLPSQ DCSQSYLELQ DSVQTGGNRY
 3331 TQFCGANYYT LPVFYSSMST AVVVFKSGVY NRNSQVQFSY QIACNREYN
 3381 QTFGNLKSPG WPQNYDNNLD CTIILRAPQN HSISLFFYWF QLEDSRQCMN
 3431 DFLEVRNGGS STSPLLDKYC SNLLPNPVFS QSNELYLHFH SDHSVTNNNGY
 3481 EIIWTSSAAG CGGTLLGDEG IFTNPGFPDS YPNNTHCEWT IVAPSGRPVS
 3531 VGFPLSIDS SGGCDQNYLI VFNGPDANSP PFGPLCGINT GIAPFYASSN
 3581 RVFIRFHAEY TTRLSGFEIM WSS SEQ ID NO. 2

FIG. 3-2

4/23

SUBSTITUTE SHEET (RULE 26)

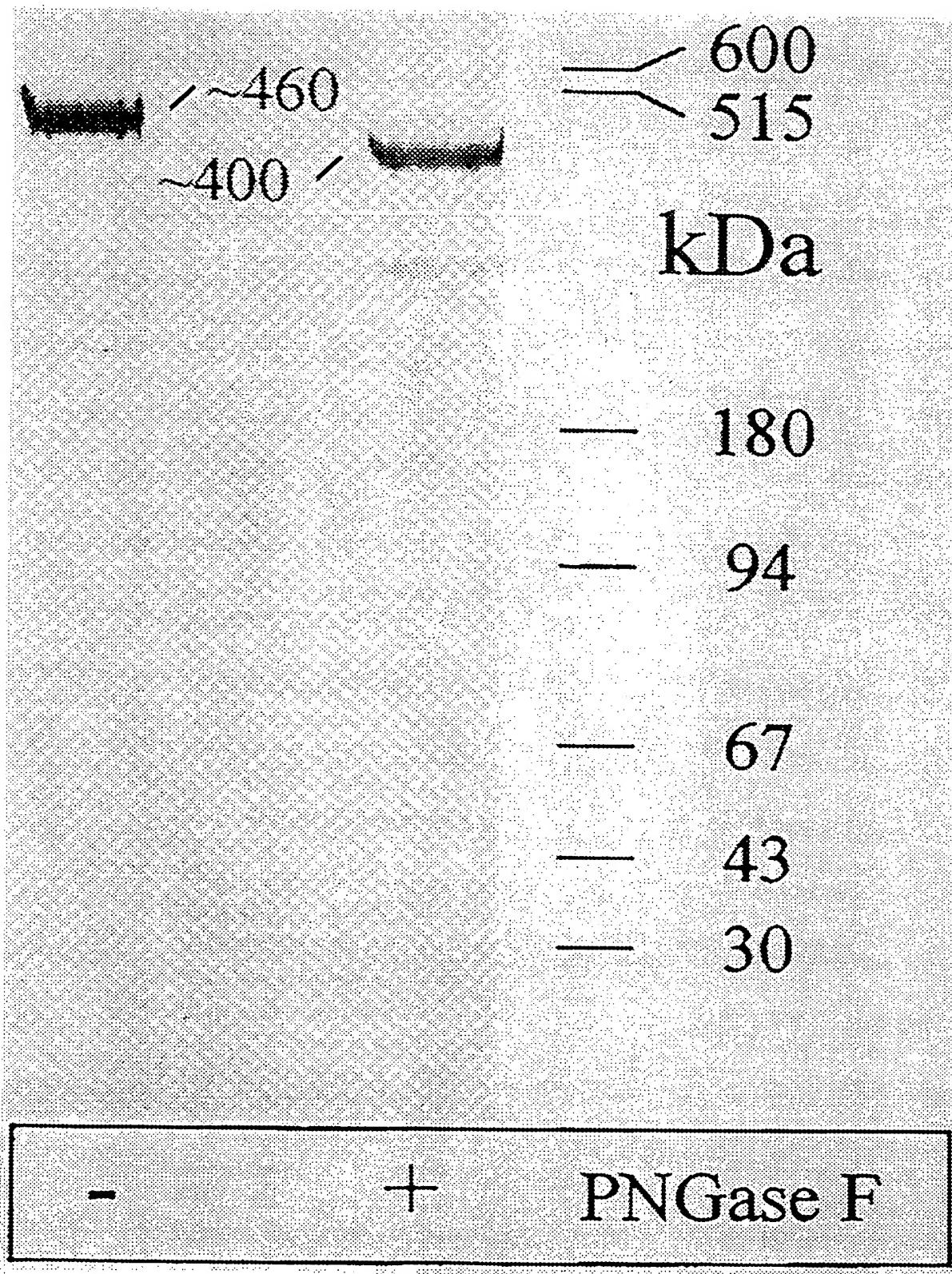


FIG. 4

5/23

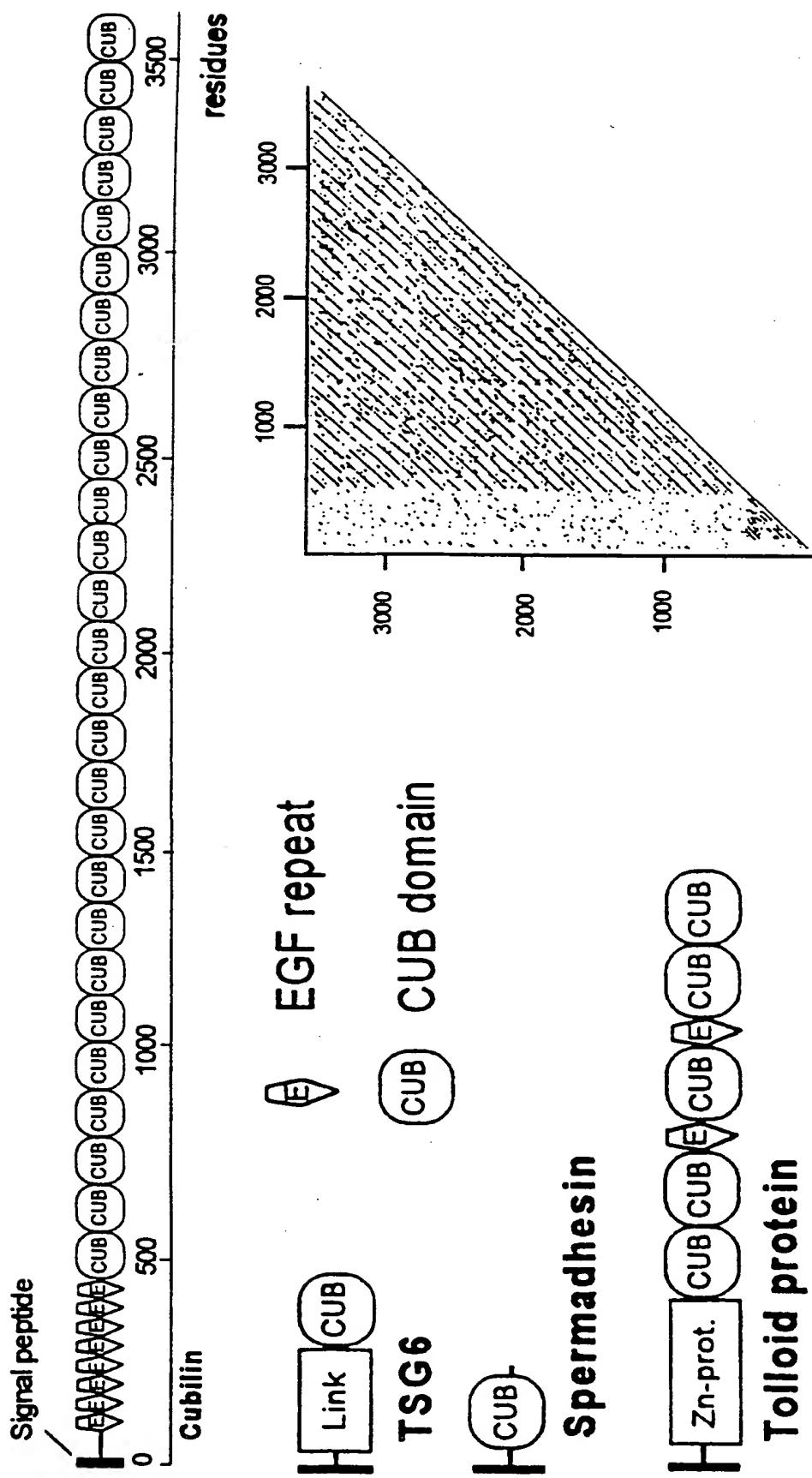


FIG. 5B

FIG. 5A

1	Cubilin-egf1	RKV . . . CS . . . SNPC . . . LNGGTCVNL . . . H . . . DSF . . . VC . . . ICPSQW . . . XGLP . . . CS
2	Cubilin-egf2	EDVNECVVYSGTPTFGCQ . . . SGSTE CVNT . . . V . . . GSF . . . RC . . . DCTPDTY . . . GPF . . . CA
3	Cubilin-egf3	SKYNDCE . . . QGSKQOLCKH . . . GICEDLQRVHHGQPNFHC . . . ICDAAGWETTPPNGI . . . SCT
4	Cubilin-egf4	EDRDEC SL . . . Q . . . PSPCSEH . . . AQCFNT . . . Q . . . GSF . . . YCGACPKGW . . . QGNGYE . . . CQ
5	Cubilin-egf5	-DINKCEII . . . N . . . NGGCSQAPLUVPCLN . . . P . . . GSF . . . SGCGNCFAEGF . . . SGDGCRY . . . CT
6	Cubilin-egf6	-PVDICSIH . . . GGC . . . HPEATCSSSPV . . . GSFLPVC . . . CPPCYTGNGYGSNGCV . . . CT
7	Cubilin-egf7	RISNICS . . . RHPC . . . VN . . . GQCIET . . . V . . . SSY . . . EC . . . KCDSCGW . . . SSEQN . . . CT
8	Cubilin-egf8	ENINDCS . . . SNPC . . . ENGGT CIDG . . . I . . . NCF . . . TC . . . DCTSSW . . . TGYY . . . CQ
9	Bmp-1-egf1	-EVDECSPRN . . . RGGCEQR . . . CLNT . . . L . . . GSY . . . KCS . . . CDPGYETAPDKFR . . . CE
10	Tolloid-egf1	-DVDECKF . . . T . . . DHGCQH . . . CINT . . . L . . . GSY . . . QCG . . . CRAGYEMQANGKT . . . CE
11	Tolloid-egf2	-DVDEC SM . . . N . . . NGGCQHR . . . CRNT . . . F . . . GSY . . . QCS . . . CRNGCYTIAENGHN . . . CT
12	C1s-egf1	-DINNECT . . . DFVDVPCSH . . . FCNNF . . . I . . . GGY . . . FCS . . . CPPEYFLKDDMKN . . . CG
13	Fibrillin-egf5	-DIDECSTIPGI . . . CE . . . GGE . . . CTNT . . . V . . . SSY . . . FC . . . KCP PGFYTSPDGT R . . . CI
14	Fibrillin-egf13	-DIDEC . . . E . . . SSP . . . CI . . . NGV . . . CKNS . . . P . . . GSE . . . IC . . . ECSEESTLDEPTKTI . . . CI
15	Fibrillin-egf26	-DVNEC LD . . . PRN . . . CR . . . SCN . . . CVNT . . . P . . . GSY . . . IC . . . DCPPDFELN PTRVG . . . CV

Residues for
 Ca^{2+} binding

$\frac{D}{N}$

$\frac{E}{N}$

CV

1	Cubilin-egf1	SEQ ID NO. 6	9	Bmp-1-egf1
2	Cubilin-egf2	SEQ ID NO. 7	10	Tolloid-egf1
3	Cubilin-egf3	SEQ ID NO. 8	11	Tolloid-egf2
4	Cubilin-egf4	SEQ ID NO. 9	12	C1a-egf1
5	Cubilin-egf5	SEQ ID NO. 10	13	Fibrillin-egf5
6	Cubilin-egf6	SEQ ID NO. 11	14	Fibrillin-egf13
7	Cubilin-egf7	SEQ ID NO. 12	15	Fibrillin-egf26
8	Cubilin-egf8	SEQ ID NO. 13		

1	Cubilin-CUB2	CGGIL...	TDNIGSITSPGYPGCN.	YPF.	GRDCVWQVLVNPNSLITF	GTLSLE	SH.
2	Cubilin-CUB5	CGEVIL...	TASTGILIESSPGEPNV.	YPR.	GVNCCTW	WV	FH.
3	Cubilin-CUB6	C...	LYDIXENDNFGMLSSPNEPN.	NYPS.	NWE	CIYRITV	DYFG
4	Cubilin-CUB9	CGG...	EMSGTAGSFSSPGYNNS.	YPH.	NKE	CIWNIRVAP	YETS
5	Cubilin-CUB12	CGGSEF...	YTLDGIFENSPDYPA.	DYEP.	NAE	CVWNIASSP	NSLN
6	Cubilin-CUB17	CGCTV...	SGDSCVIESICYP	TLPYAN.	NVF	CQWFIFRGLP	SSPG
7	Cubilin-CUB20	CCE...	IRTGDDNGVISSPNYPNL.	YSA.	WTH	CSWLLKAP	AHPT
8	Bmp-1-CUB1	CGETEQDSTGNF...	SSPEYPN.	GYSA.	HMH	CWWRISVTPGEKLTNET.	SLD.
9	Bmp-1-CUB2	CGGDV...	KDYGHTQSPNYP.	DDYRP.	SKV	CIVWRIOVSEG	LYRS
10	Tolloid-CUB2	CGGDE...	IDSPNYP.	MDYME.	DKE	CVWRITAPDNEQVAK	REDS
11	Tolloid-CUB3	CGGYV...	DATKNSGSLYSPSYPDV.	YPN.	SKQ	CWWEVVAAPPN	KEDG
12	Tolloid-CUB4	C...	KF.EITTSYGVLQSPNYP.	EDYPR.	NIY	CYWHFQTVLGHRIQ	SHQE
13	Uvs-2-CUB2	CGGAFYSSSPKT...	FTSPNYPG.	NYT.	NTN	CTWTITAP	G
14	C1s-CUB1	...	EPTMYGEILSPNYPQA.	YPS.	EVEKSWDIEV	YGHLYFTHDIEI.	S
15	Tsg6-CUB	CGGVF...	TDPKRIKSPGFPN.	EYED.	NQIC	CYWHIRLKYCORIHL	DDPG
16	Aqn-3-CUB	CGGFLKNYS...	GWIS...	Y...	YKALT	TNCVWTIEMKPGHKII	QILPLNL.

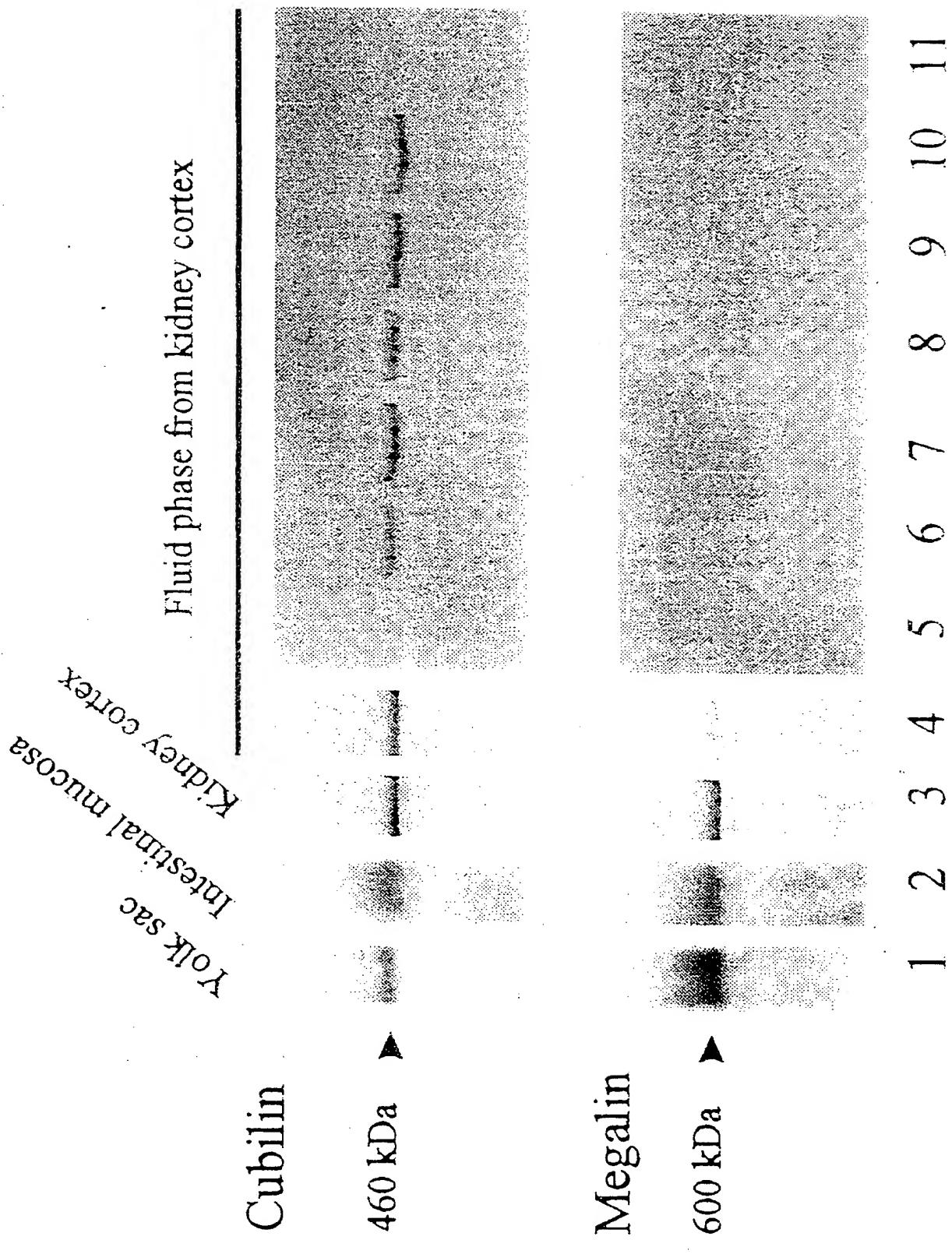
FIG. 6B-1

..continued

1 NDCSKDYLEIIRDGPPFHQD. PVLGKPFCTSL. . . STP. P. LKTTCP. . . AARIHFSSETSDK. . . GFFHITY
 2 YNCCTNDYLEIYDTA. . . AQ. TFLGRYCG. K. . . SIP. PSLTNSN. . . SIKLIFVSDSAFARE. . . GFSINY
 3 SQC. VDFEVEIIRDGGYES. PLEVGTYCG. . . SVLP. ETLLIHSN. . . KFLWLFKFSDDAETAK. . . GFSAYW
 4 . . . CNYDSLEIYAGLDENS. PREAQLCSQSPSANF. . . MQVSSTEGN. . . ELAIRFKTDSTENGR. . . GFNASW
 5 . . . CNKDFVEIREGNATGH. . . LIGRYCG. . . NSLP. GNYSSAEGH. . . SLWVRFVSDGSGTGM. . . GFQARF
 6 . . . CTKDFVEIWNHTSGR. . . VLGRYCG. . . NSTP. SSVDTSSNV. . . AS. VKFVTDGSVTEAS. . . GFLQF
 7 . . . CTSDSVTVRNGDSPGS. PVIIGRYCCQSV. . . P. RPIQSGSN. . . QLIVTFNTNNQGOTR. . . GFYATW
 8 RLCWYDYVEVRDGFWRKA. PLERGRECG. . . SKLP. EPIVSTD. . . RLWVEFRSSSNWVGR. . . GFFAVY
 9 . . . CAYDYLEEVRDGHSESS. THIGRYCG. . . YEKP. DDIKSTSS. . . RLWILKFVSDGGSINKA. . . GFAVNF
 10 . . . CAYDFVEIIRDGHSESS. RPIIGRECG. . . DKLP. PNIKTRSN. . . QMYIRFVSDSSVQKL. . . GFSARL
 11 TCKNYDYLIIYSKMRDNRLLKIGIYCG. . . HELP. PVVNSEQSI. . . TRLIEFYSDRRTVORRS. . . GFEVAKF
 12 . . . CIYDYYVATYDGRSENS. STLGIYCG. . . GREP. YAVIASTN. . . EMFMVLATDAGLQRK. . . GFEKATF
 13 ASCRIYDYYENIYNS. . . TLGAVMGPYCGPIDEH. . . SAIYSEKSN. . . SMMITMNSDFFSKQYK. . . GFSATY
 14 ENCAVDSVQIIISGDTEE. . . . GRLCCGQRSSNNPHSPIVEEFOVPYNKKQVLEKSFDFSNEERFTGFAAAYY
 15 . . . CLADYYVEIYDSYDDVH. GFVGRYCG. . . DELP. DDLISTGNV. . . MTLKELSDASVTAG. . . GFQIKY
 16 . . . TCGKEYLEVRDQRAQPDNF. . . KVCGGTGF. . . VYQSSSEHNVAT. . . VYKYSRDS. . . HEPASSFNVYF

1 Cubilin-CUB2	SEQ ID NO. 21	9 Bmp-1-CUB2	SEQ ID NO. 28
2 Cubilin-CUB5	SEQ ID NO. 22	10 Tolloid-CUB2	SEQ ID NO. 29
3 Cubilin-CUB6	SEQ ID NO. 23	11 Tolloid-CUB3	SEQ ID NO. 30
4 Cubilin-CUB9	SEQ ID NO. 24	12 Tolloid-CUB4	SEQ ID NO. 31
5 Cubilin-CUB12	SEQ ID NO. 25	13 Uvs-2-CUB2	SEQ ID NO. 32
6 Cubilin-CUB17	SEQ ID NO. 26	14 C1s-CUB1	SEQ ID NO. 33
7 Cubilin-CUB20	SEQ ID NO. 26	15 Tsg6-CUB	SEQ ID NO. 34
8 Bmp-1-CUB1	SEQ ID NO. 27	16 Aqn-3-CUB	SEQ ID NO. 35

FIG. 6B-2



10/23

SUBSTITUTE SHEET (RULE 26)

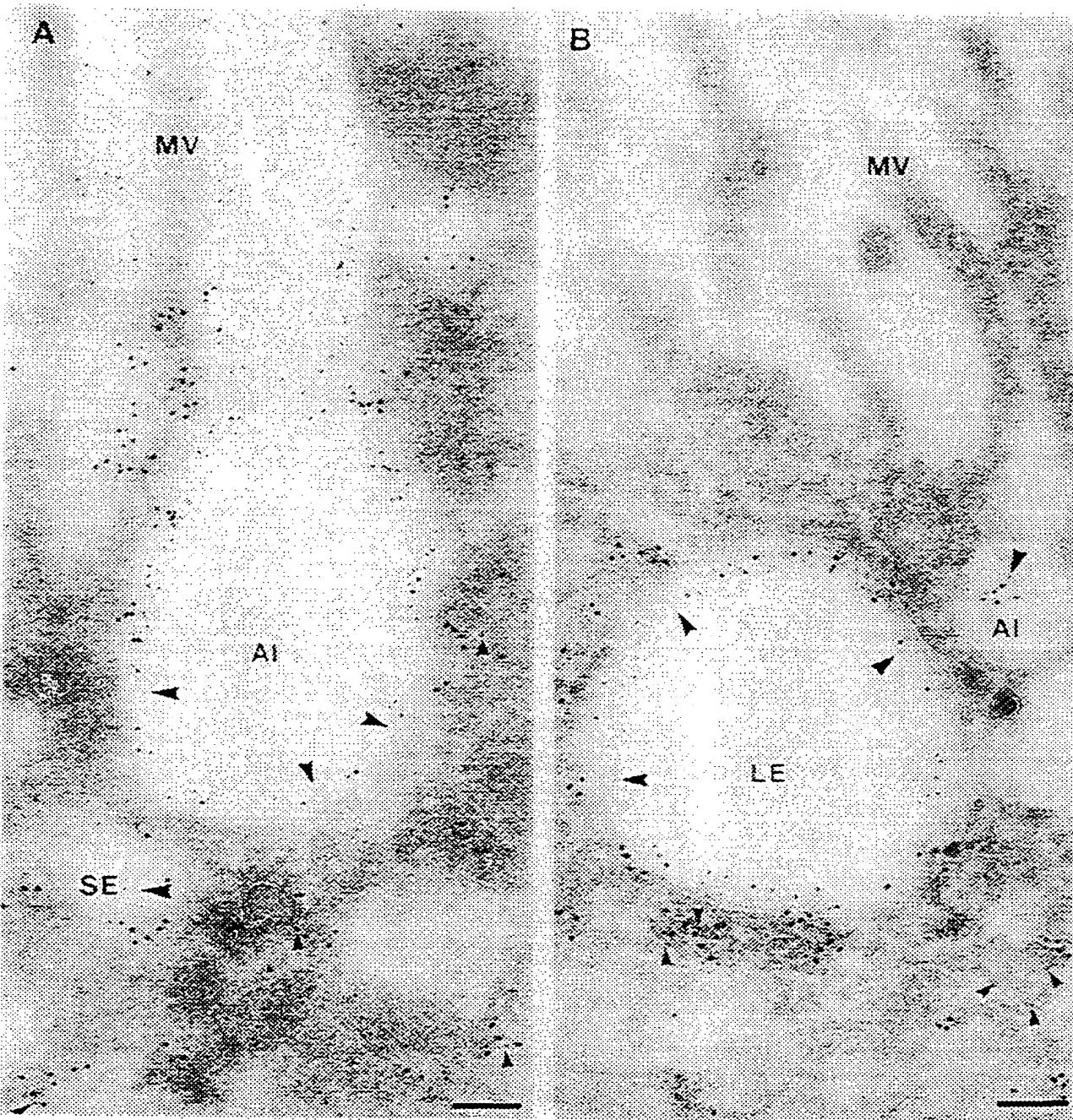


FIG. 8A

FIG. 8B

11/23

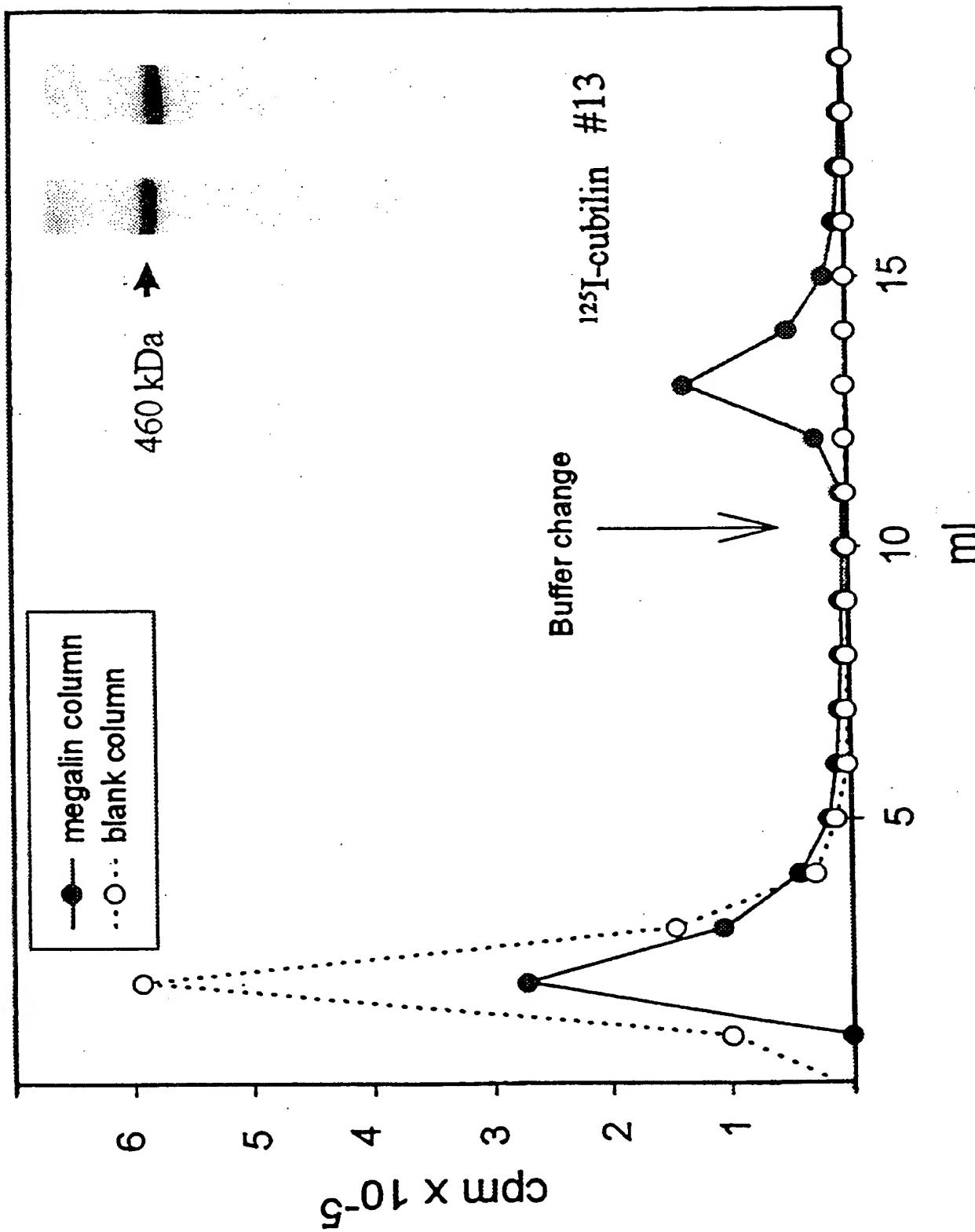


FIG. 9

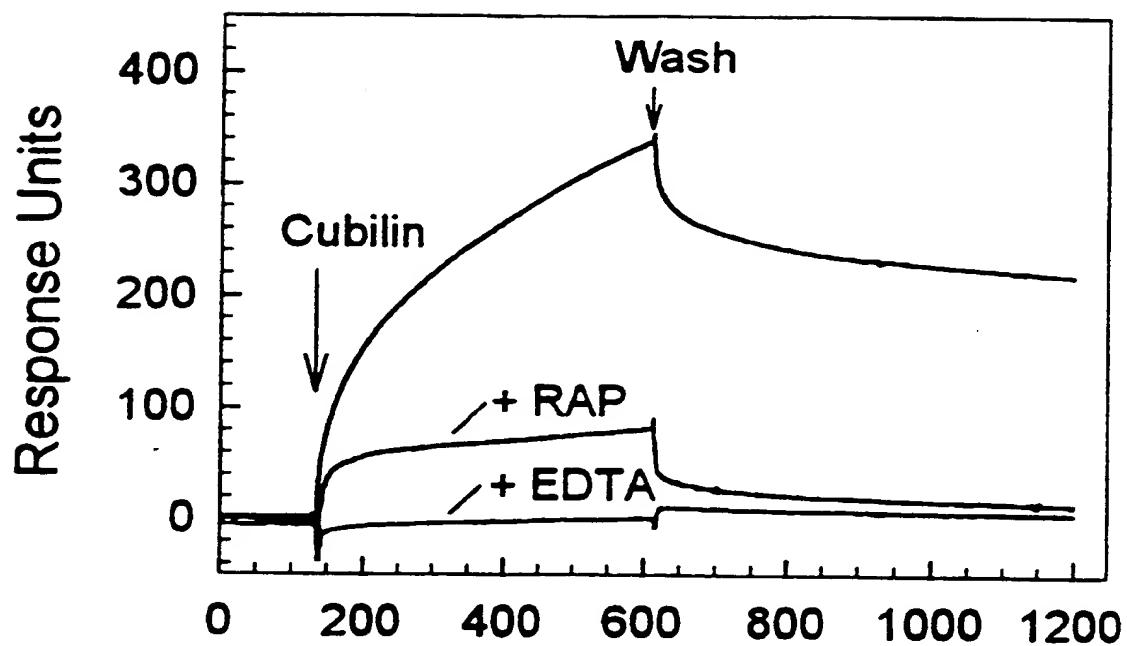


FIG. 10A

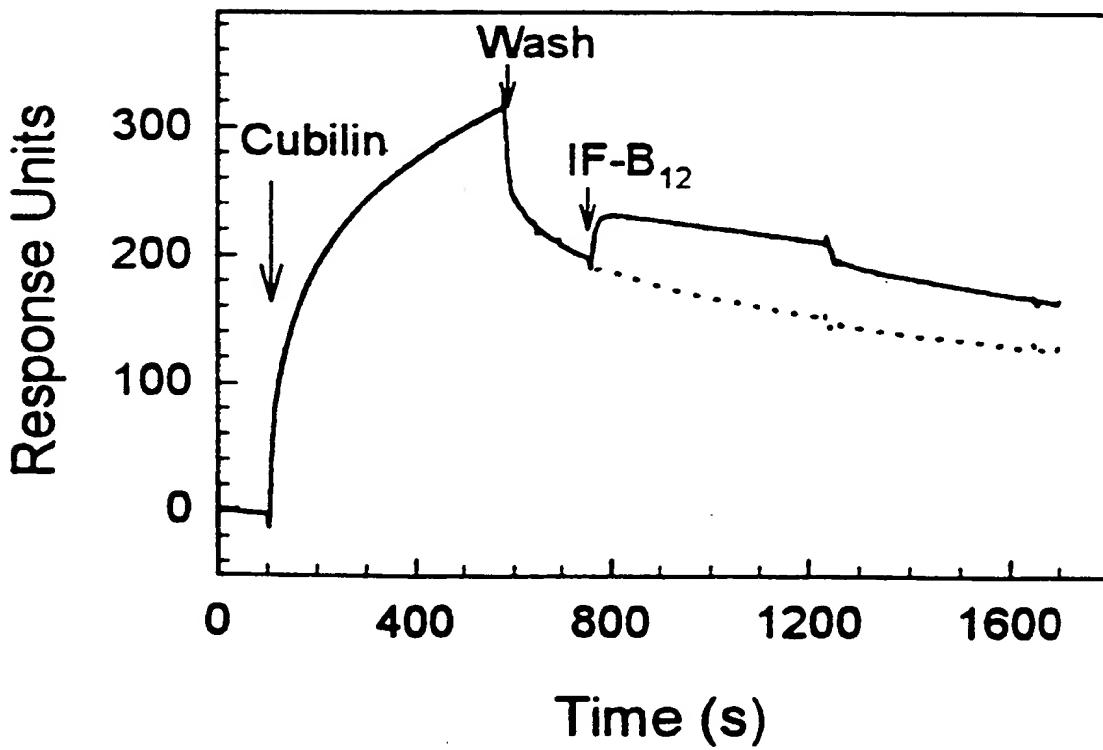


FIG. 10B

13/23

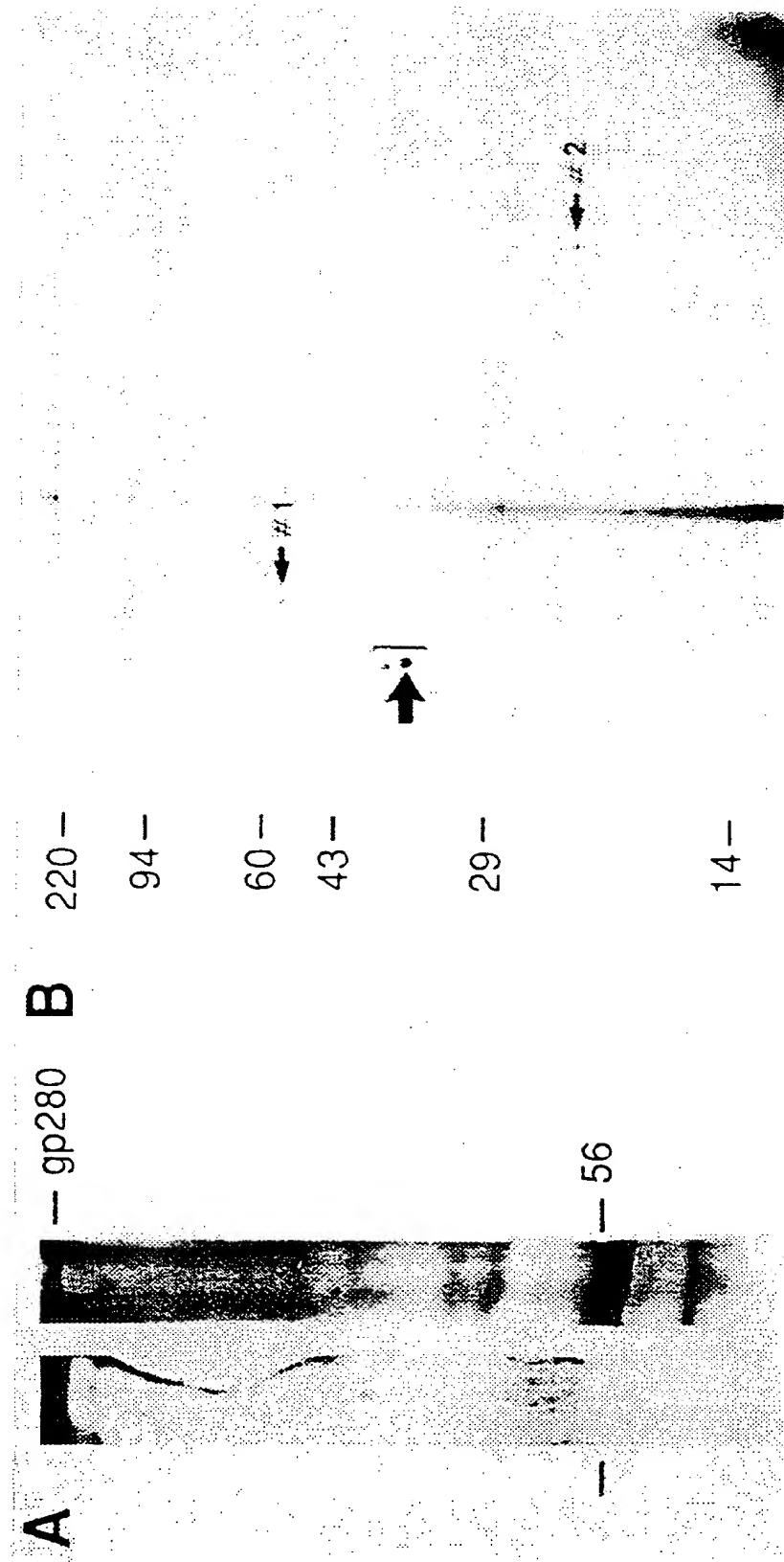


FIG. 11A

FIG. 11B

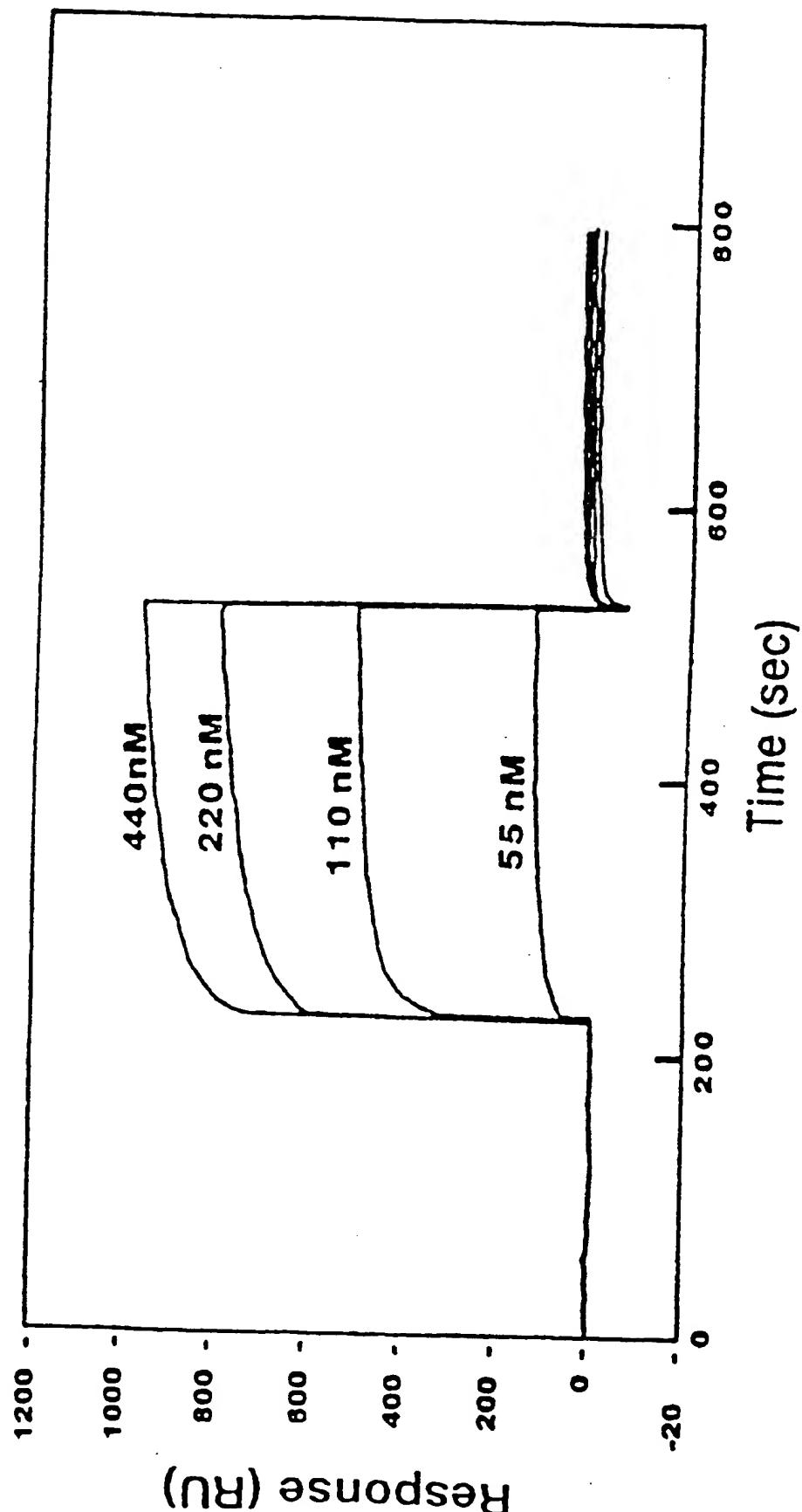


FIG. 12A

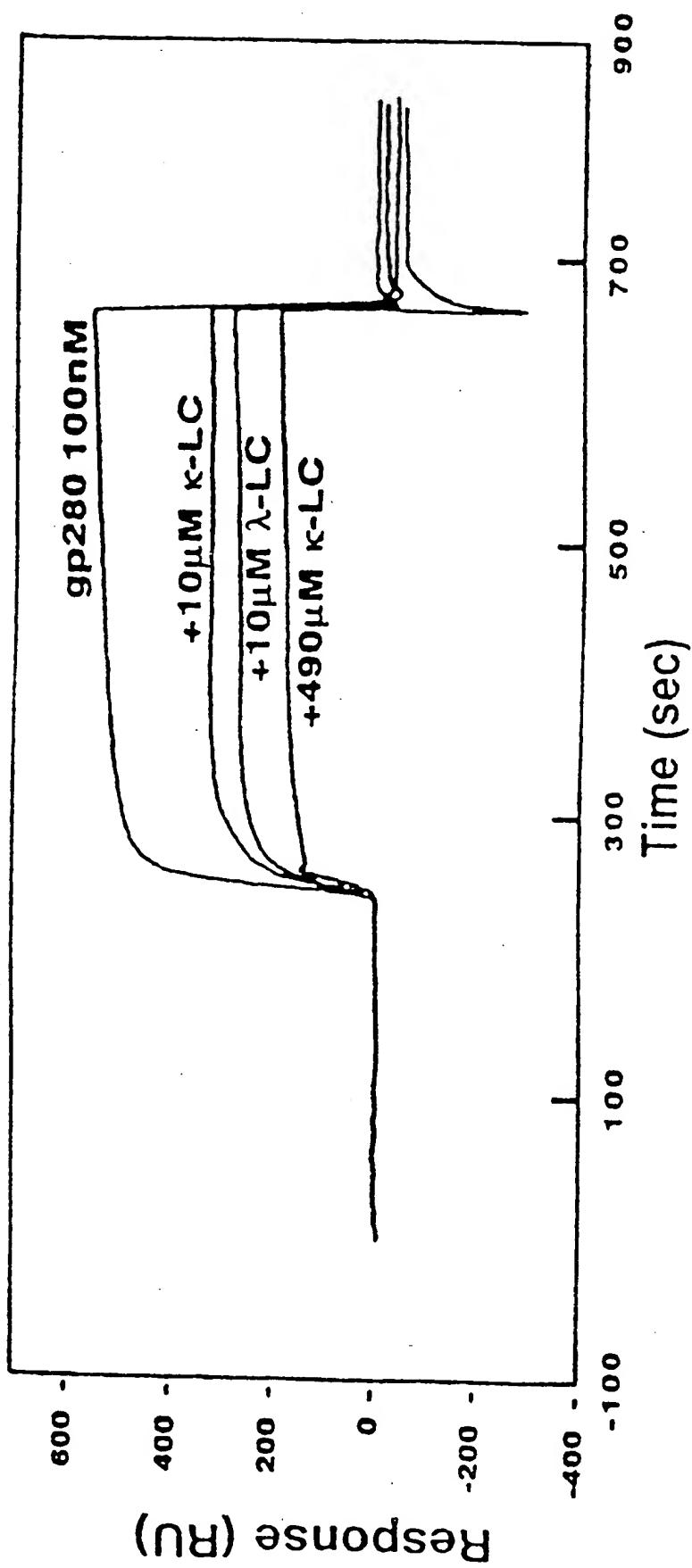


FIG. 12B

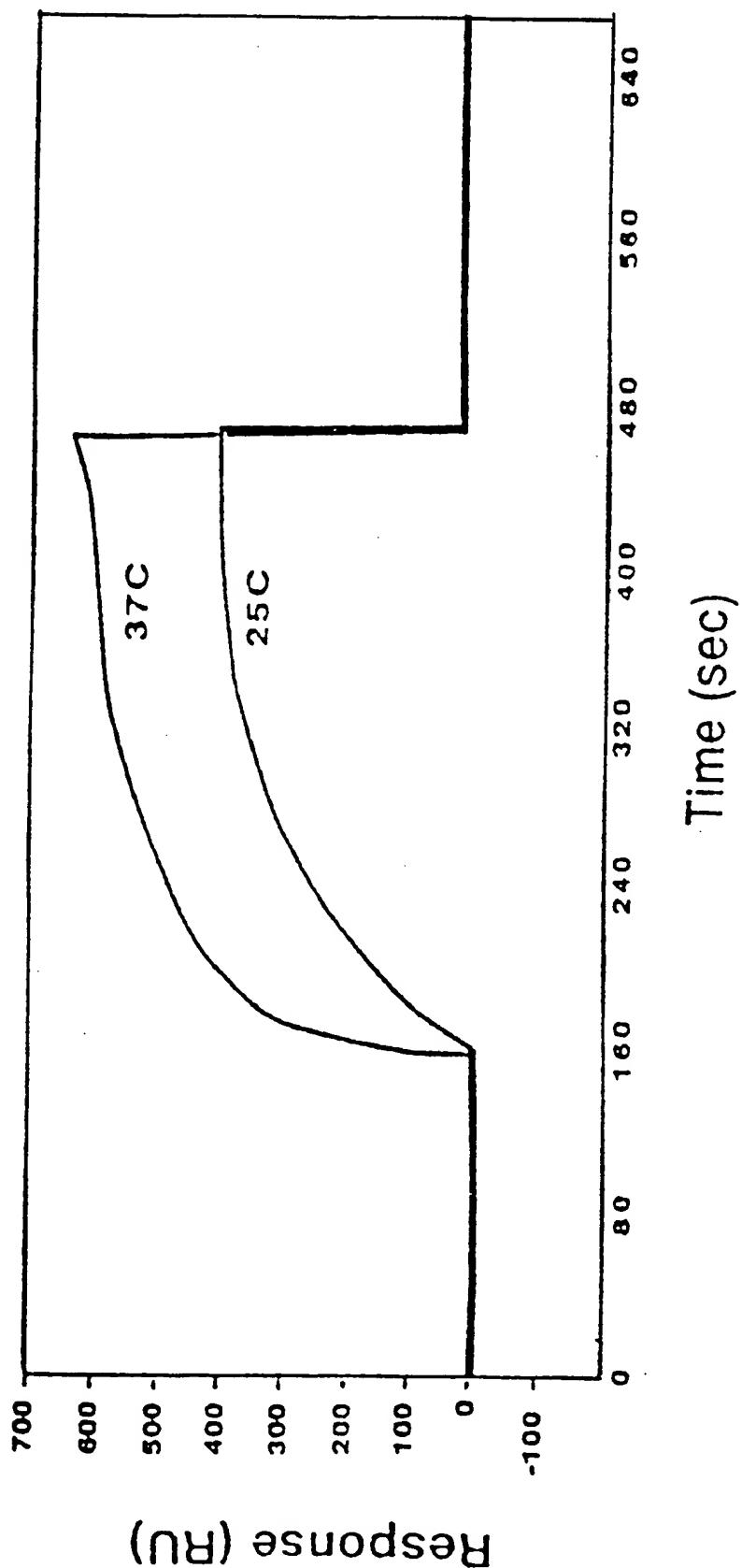


FIG. 12C

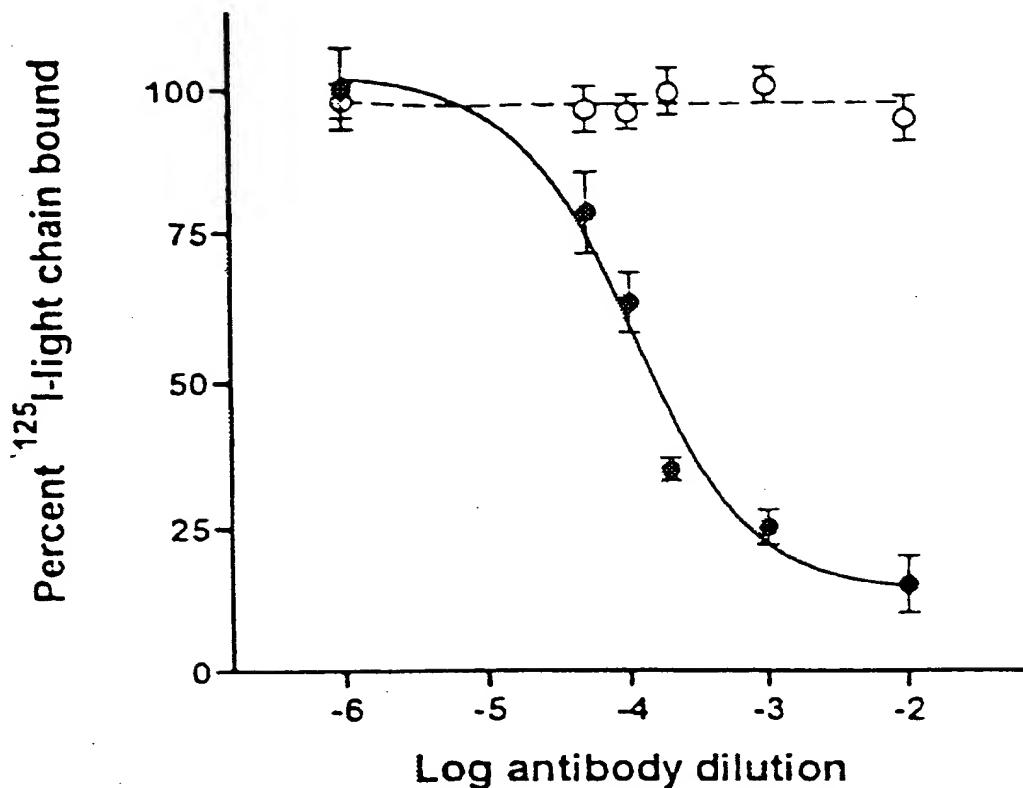


FIG. 13A

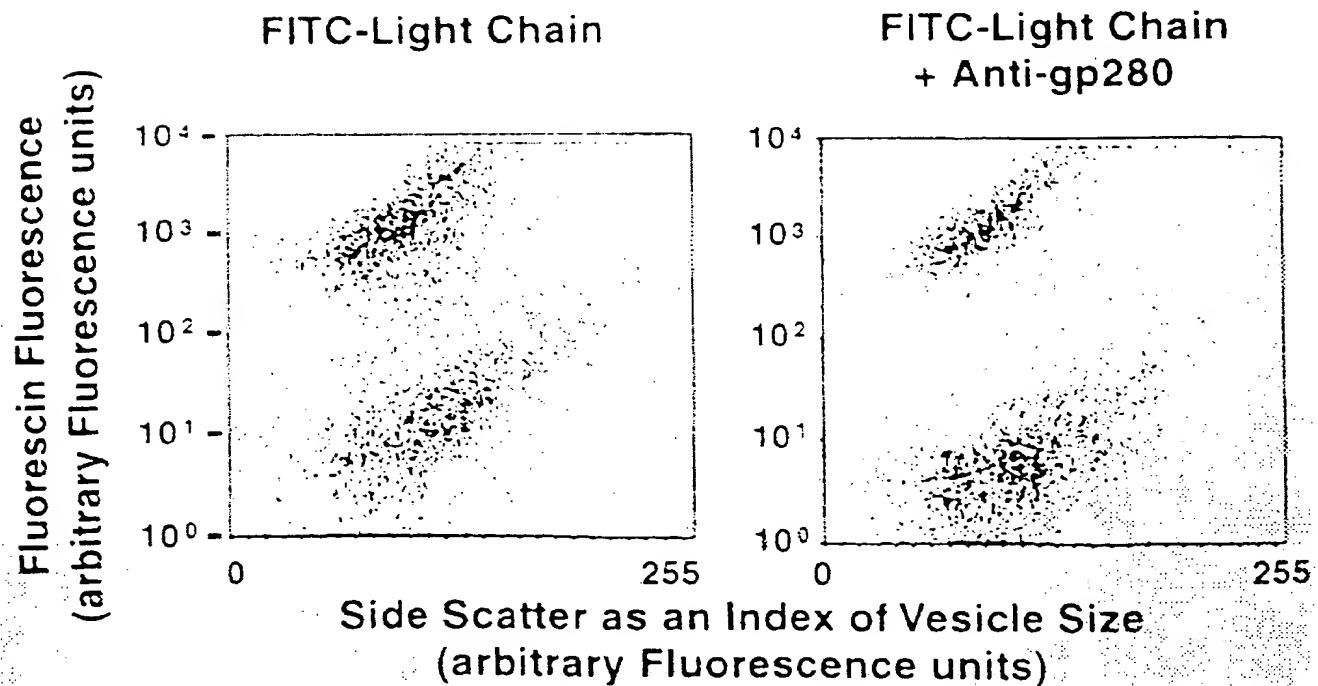


FIG. 13B

18/23

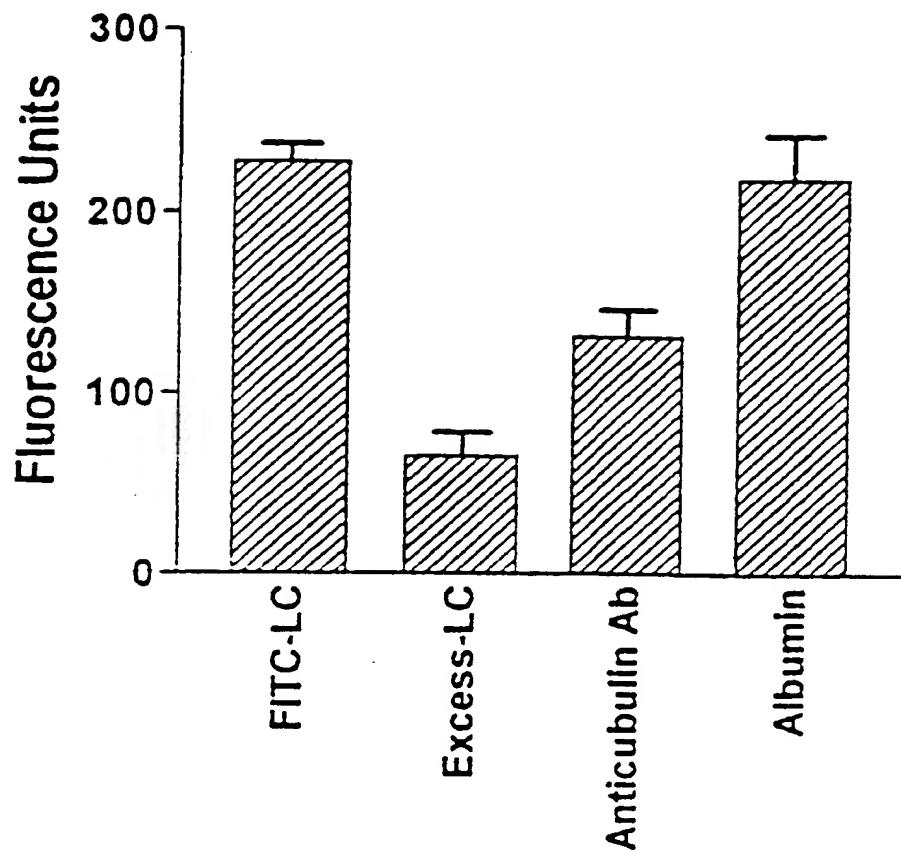


FIG. 14A

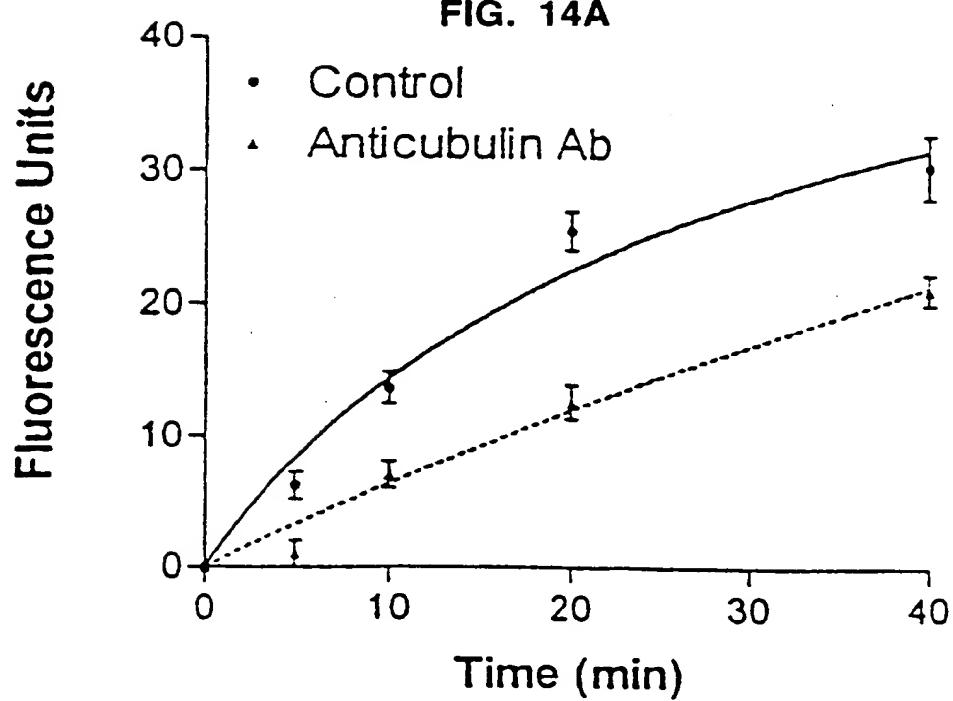


FIG. 14B

19/23

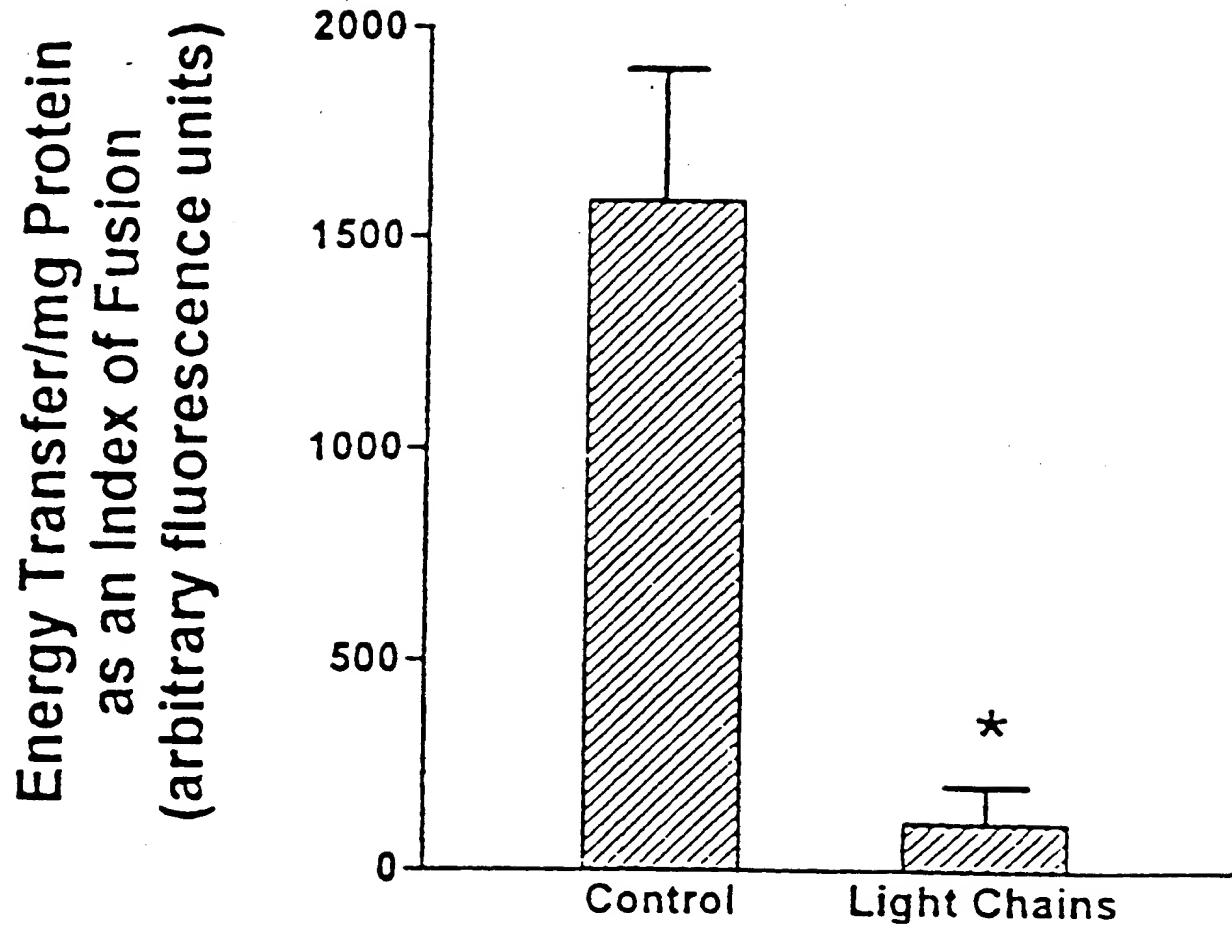


FIG. 15

20/23

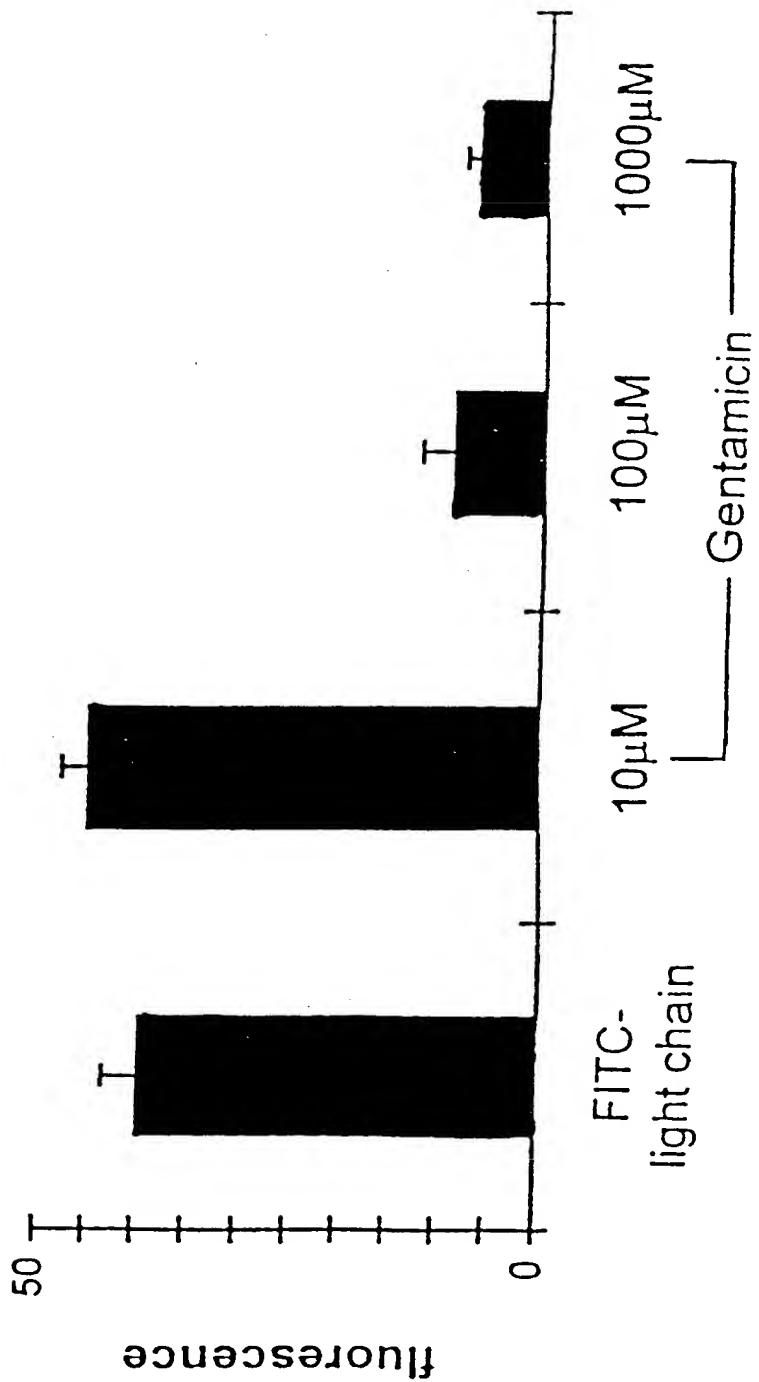


FIG. 16

21/23

SUBSTITUTE SHEET (RULE 26)

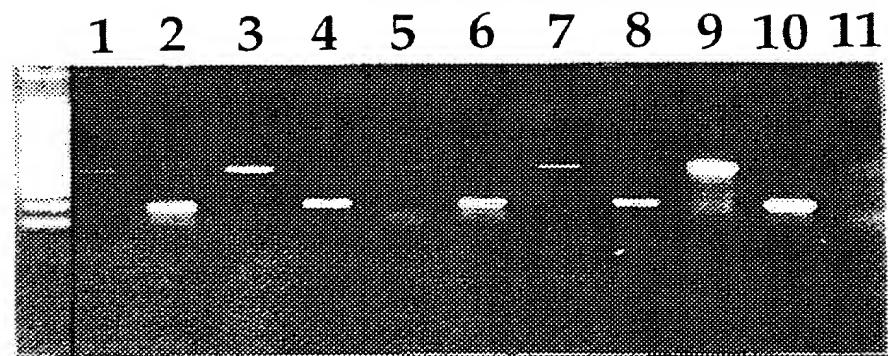


FIG. 17A

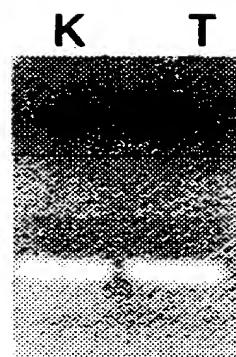


FIG. 17B

22/23

SUBSTITUTE SHEET (RULE 26)

1 2 3

FIG. 18

23/23

SUBSTITUTE SHEET (RULE 26)

SEQUENCE LISTING

<110> Hammond, Timothy G.
<120> Verroust, Pierre J.
<130> Cubilin Protein, DNA Sequences Encoding Cubilin and Uses Thereof
<140>
<141> 1999-01-21
<150> USSN 60/072,197
<151> 1998-01-22
<160> 40

<210> 1
<211> 11272
<212> DNA
<213> rat
<220>
<223> nucleic acid sequence of rat cubilin
<400> 1

atgtcctcgc agtttctctg gggtttgtt actttgttga tgatagctga attagatggc 60
aaaactggaa agccggagca gaggggacag aaaaggatcg ctgacctgca ccagcctcgg 120
atgactacgg aggagggaaa cttagtattt cttagaagct ctacccaaaa cattgaattt 180
agaactggat ccctggggaa aatcaaatta aatgatgaag accttggcga atgtttacac 240
cagatccaga gaaacaaaaga tgatattata gatttaagaa agaatacaac tggcctccct 300
caaaatatac tttagtcaagt ccaccaactt aactcgaagc tcgtggatct ggagagagat 360
tttcaaaact tacagcagaa tgtggagaga aaagtttgcg gtagcaatcc ctgccttaat 420
ggtggcacct gcgtcaacct gcacgactcc tttgtgtgt tctgccttc tcagtggaaag 480
ggtctcttct gctcagagga tgtcaacgag tgtgttagttt actcaggaac accctttggc 540
tgccagagt gatccacactg tgtgaacaca gtggggagtt tcagatgtga ctgcacgccc 600
gatacgtacg gacccatctg tgcatccaaa tataatgact gtgaacaggg ctctaaacag 660
ctctgcaagc acggcatctg tgaggattt cagcgagttc accatggaca gccgaatttt 720
cactgcacatct gtgatgctgg gtggacaaca ccgc当地atg gaatctcctg cacagaggac 780
aaagacgaat gcagcctcca gccttcacact tggcggagc atgcgcagtg tttcaataca 840
caaggctcct tctactgtgg ggcctgcccc aaagggttggc agggaaatgg ttatgaatgc 900
caggatatac atgaatgcga gatcaacaat ggaggctgtt cccaggcgcc actggccttca 960
tgcctgaaca cacctggatc attctcctgt gggactgtc cagcaggttt cagtggtgac 1020
gggagagtgt gcactcccgt ggacatctgt tcgatccaca atggaggctg ccattccagag 1080
gcaacacctgct cctcatctcc tggtaggg tccttgc acgtctgcac ctgtcctccg 1140

ggatatacag gaaatggta cgggtctaat ggatgtgtac gcctcagtaa tatttgctca 1200
aggcaccgggt gtgtaaatgg acagtgcata gaaacagtat caagtttattt ctgttaagtgt 1260
gactcaggct ggtctggaca gaactgtaca gaaaacatta atgactgttc gagcaacccc 1320
tgtttgaatg ggggcacctg tattgtatggc atcaatggct tcacctgcga ctgcacaagc 1380
tcttggactg gctattactg tcagactccc caggcagctt gcggagggat cctctcaggg 1440
acacaaggaa cctttgccta ccacagccc aatgataacct atattcacaa tgtgaactgt 1500
ttctggattg tcagaactga tgagggaaag gttctgcac tcaccttcac gtttttgat 1560
ttagaatcag caagcaattg tccccggag taccttcaga ttcatgtatgg agactcctca 1620
gcggatttcc cacttggtag atactgtggc tccagggccc cccagggat ccacagcagt 1680
gccaatgctc tctatttcca tctctactcc gagtacataa ggagtgggag aggctttaca 1740
gcaagggtggg aggcgaaagct gccagagtgt ggcggcatcc tgactgataa ttacggttct 1800
attacgtctc cggggatcccc tggaaaactac cccccaggaa gagattgtgt ctggcagggtt 1860
ttagtcaatc ctaactcctt gataacattt acctttggaa ctttgagccct ggagagccac 1920
aatgactgca gcaaagacta tttggagatt cgagacggtc ctttccacca agaccctgtt 1980
cttggaaat tctgcacttc cttgtctacc ccacccctca agactaccgg tcctgcagca 2040
agaattcatt tccattctga ctctgagacc agtgacaaag gcttccacat cacctatcta 2100
accacacagt cggatctgga ctgtgggtggg aactacacag acacggatgg cgagctcctc 2160
cttcctcctt tgtctggtcc tttcagtcac agcagacagt gtgtctatct catcacccaa 2220
gcccagggag aacaaatagt tatcaacttc acccatgtgg agctggagag ccagatggc 2280
tgttctcaca cttacatcga gttggagac catgacagct tacttcgaaa gatctgtggc 2340
aatgaaacct tgccccat tagatcagtt tctaataaag tctggatcag attgagaata 2400
gacgctttag tccagaaggc tagttcaga gctgactacc aagttgctt gggggatgt 2460
ttaagaggag aaggatttt tcgctcacct ttctatccctt acgcatatcc tggacgaaga 2520
acctgttaggt ggaccatctc ccaacccaa agacaagttg tccttctttaa cttcaactgac 2580
tttcagatcg gaagttctgc ctcctgtat acagattata ttgagattgg tccagctct 2640
gtcttggat ctcctggaaa tgagaagttt tgtagctcaa acataccgtc atttataaca 2700
tctgtataca atattcttta tgttacattt gtgaaaagtt cttccatgga aaatcgtggc 2760
ttcacggcta agttcagcag tgacaaacta gagtgtggag aagttctcac ggcacatcaca 2820
ggaattatcg aaagtccagg tcatccaaac gtctacccaa gaggtgtaaa ttgtacttgg 2880
catgtatcg tccaaacgcgg ccaactgatc cgtttggagt tcagttccctt ttacctggag 2940
tttcattaca actgcacaaa cgactacctg gaaatttatg acactgccgc tcagactttt 3000
cttggagat actgtggaaa atccatcccg cttctctta ccagcaactc taattcaata 3060
aagctgatata tttgtctga ctccgcctt gcccacgaag gctttccat aaattatgag 3120
gcaatcgatg catcatcagt atgtttat gactatacag ataattttgg gatgctctcc 3180
tccccgaact tccccatcaa ttacccctgt aactggaggt gcatctacag aatcaactgtg 3240
ggactcaacc aacagattgc attgcatttc acagacttca cttggagga ctatgggg 3300
tcacagtgtg tagatttgt agaaatcaga gacggaggct acgaaacgtc gcccgttgg 3360
gggatttact gtggctcagt tttgccttca acaatcatct cccacagtaa caagctctgg 3420

ctaaagttta agagtgacgc cgcaactcacg gcaaagggt tctcagcgta ctgggacgga 3480
tcatcaacag gctgtggagg taatctcacc acccccacagg tgctcacatc gcccaactac 3540
ccgatgcct actaccacag ctccgaatgc tactggcggc tggaagccag tcatggcagc 3600
ccttcgagc tggaattcca agacttccac ctggAACACC accccagctg ctctctggat 3660
tacttggcc gtgttgcattt cccgactacc aactcccac tgatagataa attgtgtggg 3720
gatacgacac ctgctccat ccgttccaaat aaagacgtcg tattgttaaa aactgaggaa 3780
ctgatgcaag gtcagctagg ccgtggctt gagatcaact tccggcagag atgtgacaat 3840
gtggtgatag tgaacaaaac ctttggcatc ctggagagca taaattatcc aaatccctat 3900
gataagaacc aacgttgtaa ctggaccatc caagcaacca cggcaacac cgtgaactac 3960
acgtttctgg gatttgcattt ggaaaggatc atgaactgct ccacagatta tgttagagctc 4020
tatgatggac cacaatggat gggacgctac tgtggaaata acatgcccc accagggct 4080
acaacaggct cccaaactcca cgtactgttc catacagatg ggatcaattc tggggaaaaa 4140
ggatttaaga tgcagtgggtt cactcatggc tgtggtgag agatgtctgg aaccgcaggc 4200
tccttcagca gcccggta ccccaacagc tatttcaca acaaagatg tatctggAAC 4260
attcgcgtgg ccccagggag tagcatttcg ctcaccatcc atgactttga tgtggaaat 4320
catacaagct gcaactatga ctccctggag atctatgcag gtcttgattt taactctcca 4380
agaatagccc aactgtgttc ccaatcaccg tcagcgaacc ccatgcaggt ctccagcact 4440
ggcaatgaac tagcaatccg atttaagacg gatagcactt taaatggaaag aggtttcaat 4500
gcctcgtggc gagcagtccc tggaggttgg ggtggaaatta tccagcttc cagaggagag 4560
attcattctc caaattaccc caacaactac agagctaaca cagagtgctc ctggatcatt 4620
caagttgagc gacatcaccg tttctcttg aatatcactg actttgaccc tgaagctcca 4680
gattcttgct tacgacttat ggatggctca agttccacaa acgcccgtgt cgccagtgt 4740
tgtggaaagac agcagcccc taactctatc atcgcttcag gaaacagcct ctttggaga 4800
ttccggctcg gatcttcag ccagaacagg ggctccggg ctgaatttcg ggaagagtgc 4860
ggaggccgca tcatgaccga ctcttccat actatcttct ctccactgta ccctcacaac 4920
tatctacaca accagaactg ttcttgata attgaagctc agcctccatt caatcacatt 4980
actctctcct ttactcactt tcaacttcaa aacagcacag actgtacacg ggactttgta 5040
gaaattttgg atggcaacga ctatgacgca cctgtccaag gccgttactg tggtttctcc 5100
ctgccccacc ccatcatatc atttggcaat gccttaaccg tgagggttgc cactgatcc 5160
acacgcagtt ttgagggttt ccgtccatc tattctgcatt cgacatcatc ttgtggtgga 5220
agcttctaca cacttgatgg catcttcaat agccccgact acccagcaga ctaccatcca 5280
aatgcagaaat gtgtctggaa cattgccagc tccctggca accgcctgca actgtccctc 5340
ctatccttca atttggagaa ttctctaaac tgaacaagg attttgtggaa aatccgagaa 5400
ggaaatgcca cgggccactt gattggacga tactgtggaa actccctccc tggaaattat 5460
tcgtcagctg agggacatag tctatggtc cgatttgcatt ctgatggctc aggcactggc 5520
atgggcttcc aggccaggtt caaaaatata ttggcaata ataatattgt gggaaactcat 5580
gggaaaatcg catctccctt ctggcctgga aaataccct acaactccaa ttacaaatgg 5640
gtggtaatg tggacgcata tcatattatc cacggtagaa tcttagagat ggacatagaa 5700

cccacacaacga actgctttta tgacagttta aagattttag atggatttga cactcattcc 5760
cgtctcattg gcacttactg tggtacccag acagaatcct ttagctccag tagaaactat 5820
ctgacattcc agtttcttc cgactcttct gtgtcaggaa ggggattcct tctggagtgg 5880
tttgcagtag atgtttctga tagcaccct cccaccatcg ctccaggagc ttgtggaggg 5940
tttatggtga cgggagacac tcctgtccat attttctccc cgggttggcc tagagagtat 6000
gctaatggtg ctgactgtat ctggatcatc tatgctcctg actctactgt ggaactcaac 6060
attctctcct tggacatcga accgcagcag tcatgcaatt atgacaagct gatgtaaaa 6120
gacggagaca gtgacttatac cccagagctg gctgttctgt gtggcgttaag ccctcctggg 6180
cccatccggt caactggaga atacatgtac atccgcttca cttagacac cagcgtcgcg 6240
gggacaggct tcaacgcctc ctttcacaag agctgtggg gatatttgcg tgcagatcga 6300
ggagttatca catccccaa gtatccagac acctacccctc ccaaccccaa ctgctcctgg 6360
catgttctgg tccagactgg tctgaccatc gccgtccatt ttgagcagcc tttccagatt 6420
caaaaacagag actcttttg cagtcagggg gattacttgg tgctaagaaa cggaccagat 6480
aaccattctc caccactggg accttctgga agaaatggc gtttctgcgg aatgtacgca 6540
ccgtccactc tgttcacctc aggcaatgaa atgtttgtc agttcatctc ggacagtagt 6600
aatggtggac aagggtttaa gatcagat gaggcaaaga gtttagcctg cgggggcact 6660
gtctacatcc atgatgctga ctctgacgga tacctgaccc ccccccaacta ccctgcta 6720
tatccccaaac atgccaatg catttggatc ttagaggcgc ctccaggag aagcatacag 6780
ctccaatttg aagatcaatt caatattgaa gacacacccca actgttctgt gagctatctt 6840
gaattgcgtg acggagccaa ctcgaatgca cgtctggttt ccaagttgtg tggccacact 6900
ctgcctcata gctgggtatc ctcgagagaa cgaataact tgaagttca cactgacgg 6960
ggttccagct acatgggatt caaggccaag tactctatag cttcctgtgg aggaacagtc 7020
tcaggggaca gtggagtcgat cgagagcatt ggctacccga cccttccgta tgcaaacaat 7080
gtgtttgtc agtggtttat cegaggcctc ccaggacact acctcaactc cagttttgaa 7140
gattttaaacc ttcagagctc tcctggttgt acaaaaagact ttgtggagat ctggggaaaac 7200
catacctctg gaagagttct gggagat tggaaaact ccactcctag cagtgttgac 7260
acttccagca atgttgcctc cgtcaagttt gtcacagatg gctctgtcac tgccctcagga 7320
tttaggctgc agtttaagtc cagcagacaa gtgtgtggg gggatttaca tggccctact 7380
ggcacattta cttctcccaa ctacccaaac ccaaatcctc atgcccggat ctgtgagtgg 7440
acgatcaactg tacaagaagg aaggcggatc gtcctgacgt ttaccaactt gaggctgagt 7500
acccagccat cttgtAACAG tgacccatc atcgatttca atggcattt aagcaactcg 7560
cccctactac agaaaactgtg cagccgtgtg aatgtgacca atgaattcaa atcttcagga 7620
aacaccatga aagtggattt ttctactgat ggctccggc cgtatggagg cttcaactgt 7680
tcctacaccc tcaactgaaga tgcagtgtgt ggtggatttc ttccaaagtgt ctccgggtgg 7740
aacttttctt ctcctggcta taatggaaat cgtgattatg ccagaaacct agactgtgaa 7800
tggactctca gtaatccaaa tcggggaaaat tcatccataa gtatctattt tctagaactt 7860
tccattgaaa gtcatcaaga ctgtacattt gatgtccttg agtttcgagt aggggatgct 7920
gatggggcccc tggatagagaa gttctgtacg ctgtcagcac caacagegccc ctgggtcattc 7980

ccctaccctc aggtgtggat acgcttcgtc agcaatgagc gtgtagaata tactggattc 8040
tatatacgt actcctttac agattgtggt ggaatacgga caggtgacaa tggagtgatc 8100
tcaagtccta actatccaaa cttgtacagt gcatggaccc actgttcatg gctgctgaaa 8160
gccccagaag ggcacacccat cacttcaca ctcagtgact ttcttctcga ggctcatcca 8220
acttgcactt cagactccgt cactgtcagg aatggtgact ccccaggatc gcccgtcata 8280
ggacgatact gtggacagtc agtcccaagg ccgatacagt ctggttccaa ccaacttata 8340
gtgacttttac acacaaacaa tcaagggcaa actcgtggat tttatgcaac atggaccaca 8400
aacgctttag gttgtggggg aacattccac tcagctaattgtacaatcaa atctcctcac 8460
tggcctcaga catteccaga aaacagcaga tgctcctgga cagtgatcac tcacgatagt 8520
aaacactggg agattagctt tgacagcaat ttccgaatcc ccagcagtga cagccagtgt 8580
cagaacagct tcgtgaaggt ttggggaggc aggttgatga tcaataagac cctgttagcc 8640
acgagctgtg gagatgtggc tccaagtcggc attgtcacat cagggAACat tttcactgct 8700
gtcttccaaat ctgaggagat ggcagccag ggcttctctg catccttcata tagccgggtgc 8760
ggacgcacat tcaatacctc ccctggtgac atcatctctc caaacttccc gaagcaatac 8820
gacaacaaca tgaactgcac ctacccata gacgctgacc ctcagtctct ggtcatcctg 8880
acttttgcgtt ccttcattt ggaagatcgc tcagctatca ccggAACCTG tgatcatgat 8940
ggcttgcaca tcatcaaagg tcgtaacctc tcttccactc ctctcgtgac catatgtgg 9000
tctgaaactt tgcgtccccc cactgtggac ggcccagtgt tgctcaactt ctattctgat 9060
gcatacacca cagactttgg cttcaagatttccctacagag ccattcacctg tgggtggatc 9120
tacaatgaat cctctggaaat ccttaggagc ctttcctact catacagcaa ctaccccaac 9180
aacctctact gtgtctacag cttccatgtt agaaggcagca gagtgataat aatttaggttc 9240
aatgatttcg atgtggctcc ttccaaacctt tgcacatg acttcctgga ggtgtttgat 9300
ggtcccagca ttggaaatcg atctcttggaa aagttctgtg gttccacgatc tccacaaact 9360
gttaagagca ccaatagcag cctgaccctg ctgttcaaga cagattctc tcaaacagca 9420
agagggttggaa aaatattttt ccgggagaca atagggccac agcaggatg tgggtggatac 9480
ctgaccgagg acaaccagag ctttgcgtct cctgattctg attcaatgg acgctatgac 9540
aagggtctca gctgcataatg gtacatagtt gcacctgaaa acaaactggt taagctcacc 9600
ttcaatgtgt tcactctggaa gggaccatcg tcagctggga gctgcgtcta tgattatgtg 9660
cagatagcag atggcgcaag cataaactca tatttaggtg gaaaattctg tggctccctg 9720
atgcctgccc catttatctc ttccggctac ttcccttacgt ttcaatgtgt ctctgacgta 9780
actgttgaaa tgaggggatt taatgcaaca tataccttggacatgatc tggcgggggaa 9840
acatataacg caacctcgac acctcaaaat gcgtcatcac ctcatttatac caacatcgga 9900
cggccataact ccacctgtac ttgggtcatc gcagctcccc cacagcagca ggttcagata 9960
actgtgtggg acttacagct gcccctacaa gactgctcac aaagctactt agaacttcag 10020
gattcagtagc agactgggtgg aaaccgggtg actcagttct gcggtgcgaa ctatacaacc 10080
ttgccagtg tctactcctc aatgagact gcggttgcg tttcaagtc tggagttata 10140
aacagaaaact cgcaagtgcata attctcctat cagattgcag attgcaacag agaatacaac 10200
caaacgttttgc gcaatctgaa gagtccctggg tggcctcaga actatgacaa taacctggac 10260

tgcaccatca ttctcagagc cccgcagaac cacagcattt ccctctttt ctattggttt 10320
 cagctggaaatccaaagaca atgcataat gatttcttgg aggttaagaaa cggcggcagc 10380
 agcacctcac cactgcttga caagtactgt agcaacctgc tgcccaaccc ggtcttctct 10440
 cagagcaacg aactgtatct gcactttcac agcgaccact cagtcacccaa caatggctat 10500
 gaaatttattt ggacccctc tgcgtgttgc tgggggggatctttttggg cgacgaaggg 10560
 atattcacca accctggctt tcctgacagt tacccgaaca acactcattt gaaatggacc 10620
 atttgtgctc cttctggaaag gcctgtctct gtcgggtttc ctttctcag tatcgactct 10680
 tctgggtggct gtgaccagaa ctacccata gtcttaatg gtccagacgc caactcccc 10740
 ccctttggac cgttgtgtgg catcaacact gggatagcgc ctttctatgc ttcatcaa 10800
 cgggtcttta taagggttca tgcgtgatcat acgacacggtc ttccagggtt tggaaataatg 10860
 tggagcagct gaatcaggag agctgtgtat aaccccaagga ctctgggtct gcccagtgt 10920
 gtcagatgca accctggccca gtactcattt tgcgtgttctc atccccgtct ctgcctgcgt 10980
 catcaaaacat ggactgagct tctacagcct tgaccagaga aagtgcacga ctgacttcat 11040
 acattaggcg ttctgaaggc tttgatctac tcagctgtgc acaccatctt tactggatgc 11100
 cctgttttagc cgacatcttc taaaatgttc tttaaagggtt ggaagttatc ctggcacatt 11160
 gatgtacggt tttaaaactt ggtgatacaa atggatgtat tggccacta caaaagtcaa 11220
 agtgcaggta ataaacattt ttcacatgtt aaaaaaaaaa aaaaaaaaaa aa 11272

<210>	2
<211>	3623
<212>	PRT
<213>	rat
<220>	
<223>	amino acid sequence of rat cubilin protein
<400>	2

Met	Ser	Ser	Gln	Phe	Leu	Trp	Gly	Phe	Val	Thr	Leu	Leu	Met	Ile
				5				10					15	
Ala	Glu	Leu	Asp	Gly	Lys	Thr	Gly	Lys	Pro	Glu	Gln	Arg	Gly	Gln
					20				25			30		
Lys	Arg	Ile	Ala	Asp	Leu	Gly	Gln	Pro	Arg	Met	Thr	Thr	Glu	Glu
					35				40			45		
Gly	Asn	Leu	Val	Phe	Leu	Thr	Ser	Ser	Thr	Gln	Asn	Ile	Glu	Phe
				50					55			60		
Arg	Thr	Gly	Ser	Leu	Gly	Lys	Ile	Lys	Leu	Asn	Asp	Glu	Asp	Leu
				65					70			75		
Gly	Glu	Cys	Leu	Gly	Gln	Ile	Gln	Arg	Asn	Lys	Asp	Asp	Ile	Ile
				80					85			90		
Asp	Leu	Arg	Lys	Asn	Thr	Thr	Gly	Leu	Pro	Gln	Asn	Ile	Leu	Ser

95	100	105
Gln Val Gly Gln	Leu Asn Ser Lys	Leu Val Asp Leu Glu Arg Asp
110	115	120
Phe Gln Asn Leu	Gln Gln Asn Val Glu Arg Lys Val Cys Ser Ser	
125	130	135
Asn Pro Cys Leu Asn Gly Gly Thr Cys Val Asn Leu Gly Asp Ser		
140	145	150
Phe Val Cys Ile Cys Pro Ser Gln Trp Lys Gly Leu Phe Cys Ser		
155	160	165
Glu Asp Val Asn Glu Cys Val Val Tyr Ser Gly Thr Pro Phe Gly		
170	175	180
Cys Gln Ser Gly Ser Thr Cys Val Asn Thr Val Gly Ser Phe Arg		
185	190	195
Cys Asp Cys Thr Pro Asp Thr Tyr Gly Pro Gln Cys Ala Ser Lys		
200	205	210
Tyr Asn Asp Cys Glu Gln Gly Ser Lys Gln Leu Cys Lys Gly Gly		
215	220	225
Ile Cys Glu Asp Leu Gln Arg Val Gly Gly Gln Pro Asn Phe		
230	235	240
Gly Cys Ile Cys Asp Ala Gly Trp Thr Thr Pro Pro Asn Gly Ile		
245	250	255
Ser Cys Thr Glu Asp Lys Asp Glu Cys Ser Leu Gln Pro Ser Pro		
260	265	270
Cys Ser Glu Gly Ala Gln Cys Phe Asn Thr Gln Gly Ser Phe Tyr		
275	280	285
Cys Gly Ala Cys Pro Lys Gly Trp Gln Gly Asn Gly Tyr Glu Cys		
290	295	300
Gln Asp Ile Asn Glu Cys Glu Ile Asn Asn Gly Gly Cys Ser Gln		
305	310	315
Ala Pro Leu Val Pro Cys Leu Asn Thr Pro Gly Ser Phe Ser Cys		
320	325	330
Gly Asn Cys Pro Ala Gly Phe Ser Gly Asp Gly Arg Val Cys Thr		
335	340	345
Pro Val Asp Ile Cys Ser Ile Gly Asn Gly Gly Cys Gly Pro Glu		
350	355	360
Ala Thr Cys Ser Ser Ser Pro Val Leu Gly Ser Phe Leu Pro Val		
365	370	375
Cys Thr Cys Pro Pro Gly Tyr Thr Gly Asn Gly Tyr Gly Ser Asn		

380	385	390
Gly Cys Val Arg Leu Ser Asn Ile Cys	Ser Arg Gly Pro Cys Val	
395	400	405
Asn Gly Gln Cys Ile Glu Thr Val Ser	Ser Tyr Phe Cys Lys Cys	
410	415	420
Asp Ser Gly Trp Ser Gly Gln Asn Cys	Thr Glu Asn Ile Asn Asp	
425	430	435
Cys Ser Ser Asn Pro Cys Leu Asn Gly	Gly Thr Cys Ile Asp Gly	
440	445	450
Ile Asn Gly Phe Thr Cys Asp Cys Thr	Ser Ser Trp Thr Gly Tyr	
455	460	465
Tyr Cys Gln Thr Pro Gln Ala Ala Cys	Gly Gly Ile Leu Ser Gly	
470	475	480
Thr Gln Gly Thr Phe Ala Tyr Gly Ser	Pro Asn Asp Thr Tyr Ile	
485	490	495
Gly Asn Val Asn Cys Phe Trp Ile Val Arg	Thr Asp Glu Glu Lys	
500	505	510
Val Leu Gly Val Thr Phe Thr Phe Asp	Leu Glu Ser Ala Ser	
515	520	525
Asn Cys Pro Arg Glu Tyr Leu Gln Ile	Gly Asp Gly Asp Ser Ser	
530	535	540
Ala Asp Phe Pro Leu Gly Arg Tyr Cys	Gly Ser Arg Pro Pro Gln	
545	550	555
Gly Ile Gly Ser Ser Ala Asn Ala Leu	Tyr Phe Gly Leu Tyr Ser	
560	565	570
Glu Tyr Ile Arg Ser Gly Arg Gly Phe	Thr Ala Arg Trp Glu Ala	
575	580	585
Lys Leu Pro Glu Cys Gly Gly Ile Leu	Thr Asp Asn Tyr Gly Ser	
590	595	600
Ile Thr Ser Pro Gly Tyr Pro Gly Asn	Tyr Pro Pro Gly Arg Asp	
605	610	615
Cys Val Trp Gln Val Leu Val Asn Pro	Asn Ser Leu Ile Thr Phe	
620	625	630
Thr Phe Gly Thr Leu Ser Leu Glu Ser	Gly Asn Asp Cys Ser Lys	
635	640	645
Asp Tyr Leu Glu Ile Arg Asp Gly Pro	Phe Gly Gln Asp Pro Val	
650	655	660
Leu Gly Lys Phe Cys Thr Ser Leu Ser	Thr Pro Pro Leu Lys Thr	

665	670	675
Thr Gly Pro Ala Ala Arg Ile Gly Phe	Gly Ser Asp Ser Glu	Thr
680	685	690
Ser Asp Lys Gly Phe Gly Ile Thr Tyr	Leu Thr Thr Gln Ser Asp	
695	700	705
Leu Asp Cys Gly Gly Asn Tyr Thr Asp	Thr Asp Gly Glu Leu Leu	
710	715	720
Leu Pro Pro Leu Ser Gly Pro Phe Ser	Gly Ser Arg Gln Cys Val	
725	730	735
Tyr Leu Ile Thr Gln Ala Gln Gly Glu	Gln Ile Val Ile Asn Phe	
740	745	750
Thr Gly Val Glu Leu Glu Ser Gln Met	Gly Cys Ser Gly Thr Tyr	
755	760	765
Ile Glu Val Gly Asp Gly Asp Ser Leu	Leu Arg Lys Ile Cys Gly	
770	775	780
Asn Glu Thr Leu Phe Pro Ile Arg Ser	Val Ser Asn Lys Val Trp	
785	790	795
Ile Arg Leu Arg Ile Asp Ala Leu Val	Gln Lys Ala Ser Phe Arg	
800	805	810
Ala Asp Tyr Gln Val Ala Cys Gly Gly	Met Leu Arg Gly Glu Gly	
815	820	825
Phe Phe Arg Ser Pro Phe Tyr Pro Asn Ala	Tyr Pro Gly Arg Arg	
830	835	840
Thr Cys Arg Trp Thr Ile Ser Gln Pro	Gln Arg Gln Val Val Leu	
845	850	855
Leu Asn Phe Thr Asp Phe Gln Ile Gly	Ser Ser Ala Ser Cys Asp	
860	865	870
Thr Asp Tyr Ile Glu Ile Gly Pro Ser	Ser Val Leu Gly Ser Pro	
875	880	885
Gly Asn Glu Lys Phe Cys Ser Ser Asn	Ile Pro Ser Phe Ile Thr	
890	895	900
Ser Val Tyr Asn Ile Leu Tyr Val Thr	Phe Val Lys Ser Ser Ser	
905	910	915
Met Glu Asn Arg Gly Phe Thr Ala Lys	Phe Ser Ser Asp Lys Leu	
920	925	930
Glu Cys Gly Glu Val Leu Thr Ala Ser	Thr Gly Ile Ile Glu Ser	
935	940	945
Pro Gly Gly Pro Asn Val Tyr Pro Arg	Gly Val Asn Cys Thr Trp	

950	955	960
Gly Val Val Val Gln Arg Gly Gln Leu Ile Arg Leu Glu Phe Ser		
965	970	975
Ser Phe Tyr Leu Glu Phe Gly Tyr Asn Cys Thr Asn Asp Tyr Leu		
980	985	990
Glu Ile Tyr Asp Thr Ala Ala Gln Thr Phe Leu Gly Arg Tyr Cys		
995	1000	1005
Gly Lys Ser Ile Pro Pro Ser Leu Thr Ser Asn Ser Asn Ser Ile		
1010	1015	1020
Lys Leu Ile Phe Val Ser Asp Ser Ala Leu Ala Gly Glu Gly Phe		
1025	1030	1035
Ser Ile Asn Tyr Glu Ala Ile Asp Ala Ser Ser Val Cys Leu Tyr		
1040	1045	1050
Asp Tyr Thr Asp Asn Phe Gly Met Leu Ser Ser Pro Asn Phe Pro		
1055	1060	1065
Asn Asn Tyr Pro Ser Asn Trp Glu Cys Ile Tyr Arg Ile Thr Val		
1070	1075	1080
Gly Leu Asn Gln Gln Ile Ala Leu Gly Phe Thr Asp Phe Thr Leu		
1085	1090	1095
Glu Asp Tyr Phe Gly Ser Gln Cys Val Asp Phe Val Glu Ile Arg		
1100	1105	1110
Asp Gly Gly Tyr Glu Thr Ser Pro Leu Val Gly Ile Tyr Cys Gly		
1115	1120	1125
Ser Val Leu Pro Pro Thr Ile Ile Ser Gly Ser Asn Lys Leu Trp		
1130	1135	1140
Leu Lys Phe Lys Ser Asp Ala Ala Leu Thr Ala Lys Gly Phe Ser		
1145	1150	1155
Ala Tyr Trp Asp Gly Ser Ser Thr Gly Cys Gly Gly Asn Leu Thr		
1160	1165	1170
Thr Pro Gln Val Leu Thr Ser Pro Asn Tyr Pro Met Pro Tyr Tyr		
1175	1180	1185
Gly Ser Ser Glu Cys Tyr Trp Arg Leu Glu Ala Ser Gly Gly Ser		
1190	1195	1200
Pro Phe Glu Leu Glu Phe Gln Asp Phe Gly Leu Glu Gly Gly Pro		
1205	1210	1215
Ser Cys Ser Leu Asp Tyr Leu Gly Arg Val Asp Gly Pro Thr Thr		
1220	1225	1230
Asn Ser Arg Leu Ile Asp Lys Leu Cys Gly Asp Thr Thr Pro Ala		

1235	1240	1245
Pro Ile Arg Ser Asn Lys Asp Val Val	Leu Leu Lys Thr Glu Glu	
1250	1255	1260
Leu Met Gln Gly Gln Leu Gly Arg Gly	Phe Glu Ile Asn Phe Arg	
1265	1270	1275
Gln Arg Cys Asp Asn Val Val Ile Val	Asn Lys Thr Phe Gly Ile	
1280	1285	1290
Leu Glu Ser Ile Asn Tyr Pro Asn Pro	Tyr Asp Lys Asn Gln Arg	
1295	1300	1305
Cys Asn Trp Thr Ile Gln Ala Thr Thr	Gly Asn Thr Val Asn Tyr	
1310	1315	1320
Thr Phe Leu Gly Phe Asp Val Glu Ser	Tyr Met Asn Cys Ser Thr	
1325	1330	1335
Asp Tyr Val Glu Leu Tyr Asp Gly Pro	Gln Trp Met Gly Arg Tyr	
1340	1345	1350
Cys Gly Asn Asn Met Pro Pro Pro Gly	Ala Thr Thr Gly Ser Gln	
1355	1360	1365
Leu Gly Val Leu Phe Gly Thr Asp Gly	Ile Asn Ser Gly Glu Lys	
1370	1375	1380
Gly Phe Lys Met Gln Trp Phe Thr Gly	Gly Cys Gly Gly Glu Met	
1385	1390	1395
Ser Gly Thr Ala Gly Ser Phe Ser Ser	Pro Gly Tyr Pro Asn Ser	
1400	1405	1410
Tyr Pro Gly Asn Lys Glu Cys Ile Trp	Asn Ile Arg Val Ala Pro	
1415	1420	1425
Gly Ser Ser Ile Gln Leu Thr Ile Gly	Asp Phe Asp Val Glu Tyr	
1430	1435	1440
Gly Thr Ser Cys Asn Tyr Asp Ser Leu	Glu Ile Tyr Ala Gly Leu	
1445	1450	1455
Asp Phe Asn Ser Pro Arg Ile Ala Gln	Leu Cys Ser Gln Ser Pro	
1460	1465	1470
Ser Ala Asn Pro Met Gln Val Ser Ser	Thr Gly Asn Glu Leu Ala	
1475	1480	1485
Ile Arg Phe Lys Thr Asp Ser Thr Leu	Asn Gly Arg Gly Phe Asn	
1490	1495	1500
Ala Ser Trp Arg Ala Val Pro Gly Gly	Cys Gly Gly Ile Ile Gln	
1505	1510	1515
Leu Ser Arg Gly Glu Ile Gly Ser Pro	Asn Tyr Pro Asn Asn Tyr	

1520	1525	1530
Arg Ala Asn Thr	Glu Cys Ser Trp Ile Ile Gln Val Glu Arg Gly	
1535	1540	1545
Gly Arg Val Leu	Leu Asn Ile Thr Asp Phe Asp Leu Glu Ala Pro	
1550	1555	1560
Asp Ser Cys Leu	Arg Leu Met Asp Gly Ser Ser Ser Thr Asn Ala	
1565	1570	1575
Arg Val Ala Ser	Val Cys Gly Arg Gln Gln Pro Pro Asn Ser Ile	
1580	1585	1590
Ile Ala Ser Gly	Asn Ser Leu Phe Val Arg Phe Arg Ser Gly Ser	
1595	1600	1605
Ser Ser Gln Asn	Arg Gly Phe Arg Ala Glu Phe Arg Glu Glu Cys	
1610	1615	1620
Gly Gly Arg Ile	Met Thr Asp Ser Ser Asp Thr Ile Phe Ser Pro	
1625	1630	1635
Leu Tyr Pro Gly	Asn Tyr Leu Gly Asn Gln Asn Cys Ser Trp Ile	
1640	1645	1650
Ile Glu Ala Gln	Pro Pro Phe Asn Gly Ile Thr Leu Ser Phe Thr	
1655	1660	1665
Gly Phe Gln Leu	Gln Asn Ser Thr Asp Cys Thr Arg Asp Phe Val	
1670	1675	1680
Glu Ile Leu Asp	Gly Asn Asp Tyr Asp Ala Pro Val Gln Gly Arg	
1685	1690	1695
Tyr Cys Gly Phe	Ser Leu Pro Gly Pro Ile Ile Ser Phe Gly Asn	
1700	1705	1710
Ala Leu Thr Val	Arg Phe Val Thr Asp Ser Thr Arg Ser Phe Glu	
1715	1720	1725
Gly Phe Arg Ala	Ile Tyr Ser Ala Ser Thr Ser Ser Cys Gly Gly	
1730	1735	1740
Ser Phe Tyr Thr	Leu Asp Gly Ile Phe Asn Ser Pro Asp Tyr Pro	
1745	1750	1755
Ala Asp Tyr Gly	Pro Asn Ala Glu Cys Val Trp Asn Ile Ala Ser	
1760	1765	1770
Ser Pro Gly Asn	Arg Leu Gln Leu Ser Phe Leu Ser Phe Asn Leu	
1775	1780	1785
Glu Asn Ser Leu	Asn Cys Asn Lys Asp Phe Val Glu Ile Arg Glu	
1790	1795	1800
Gly Asn Ala Thr	Gly Leu Ile Gly Arg Tyr Cys Gly Asn Ser	

1805	1810	1815
Leu Pro Gly Asn Tyr Ser Ser Ala Glu Gly Gly Ser Leu Trp Val		
1820	1825	1830
Arg Phe Val Ser Asp Gly Ser Gly Thr Gly Met Gly Phe Gln Ala		
1835	1840	1845
Arg Phe Lys Asn Ile Phe Gly Asn Asn Asn Ile Val Gly Thr Gly		
1850	1855	1860
Gly Lys Ile Ala Ser Pro Phe Trp Pro Gly Lys Tyr Pro Tyr Asn		
1865	1870	1875
Ser Asn Tyr Lys Trp Val Val Asn Val Asp Ala Tyr Gly Ile Ile		
1880	1885	1890
Gly Gly Arg Ile Leu Glu Met Asp Ile Glu Pro Thr Thr Asn Cys		
1895	1900	1905
Phe Tyr Asp Ser Leu Lys Ile Tyr Asp Gly Phe Asp Thr Gly Ser		
1910	1915	1920
Arg Leu Ile Gly Thr Tyr Cys Gly Thr Gln Thr Glu Ser Phe Ser		
1925	1930	1935
Ser Ser Arg Asn Tyr Leu Thr Phe Gln Phe Ser Ser Asp Ser Ser		
1940	1945	1950
Val Ser Gly Arg Gly Phe Leu Leu Glu Trp Phe Ala Val Asp Val		
1955	1960	1965
Ser Asp Ser Thr Pro Pro Thr Ile Ala Pro Gly Ala Cys Gly Gly		
1970	1975	1980
Phe Met Val Thr Gly Asp Thr Pro Val Gly Ile Phe Ser Pro Gly		
1985	1990	1995
Trp Pro Arg Glu Tyr Ala Asn Gly Ala Asp Cys Ile Trp Ile Ile		
2000	2005	2010
Tyr Ala Pro Asp Ser Thr Val Glu Leu Asn Ile Leu Ser Leu Asp		
2015	2020	2025
Ile Glu Pro Gln Gln Ser Cys Asn Tyr Asp Lys Leu Ile Val Lys		
2030	2035	2040
Asp Gly Asp Ser Asp Leu Ser Pro Glu Leu Ala Val Leu Cys Gly		
2045	2050	2055
Val Ser Pro Pro Gly Pro Ile Arg Ser Thr Gly Glu Tyr Met Tyr		
2060	2065	2070
Ile Arg Phe Thr Ser Asp Thr Ser Val Ala Gly Thr Gly Phe Asn		
2075	2080	2085
Ala Ser Phe Gly Lys Ser Cys Gly Gly Tyr Leu Gly Ala Asp Arg		

2090	2095	2100
Gly Val Ile Thr Ser Pro Lys Tyr Pro Asp Thr Tyr Leu Pro Asn		
2105	2110	2115
Leu Asn Cys Ser Trp Gly Val Leu Val Gln Thr Gly Leu Thr Ile		
2120	2125	2130
Ala Val Gly Phe Glu Gln Pro Phe Gln Ile Gln Asn Arg Asp Ser		
2135	2140	2145
Phe Cys Ser Gln Gly Asp Tyr Leu Val Leu Arg Asn Gly Pro Asp		
2150	2155	2160
Asn Gly Ser Pro Pro Leu Gly Pro Ser Gly Arg Asn Gly Arg Phe		
2165	2170	2175
Cys Gly Met Tyr Ala Pro Ser Thr Leu Phe Thr Ser Gly Asn Glu		
2180	2185	2190
Met Phe Val Gln Phe Ile Ser Asp Ser Ser Asn Gly Gly Gln Gly		
2195	2200	2205
Phe Lys Ile Arg Tyr Glu Ala Lys Ser Leu Ala Cys Gly Gly Thr		
2210	2215	2220
Val Tyr Ile Gly Asp Ala Asp Ser Asp Gly Tyr Leu Thr Ser Pro		
2225	2230	2235
Asn Tyr Pro Ala Asn Tyr Pro Gln Gly Ala Glu Cys Ile Trp Ile		
2240	2245	2250
Leu Glu Ala Pro Pro Gly Arg Ser Ile Gln Leu Gln Phe Glu Asp		
2255	2260	2265
Gln Phe Asn Ile Glu Asp Thr Pro Asn Cys Ser Val Ser Tyr Leu		
2270	2275	2280
Glu Leu Arg Asp Gly Ala Asn Ser Asn Ala Arg Leu Val Ser Lys		
2285	2290	2295
Leu Cys Gly Gly Thr Leu Pro Gly Ser Trp Val Ser Ser Arg Glu		
2300	2305	2310
Arg Ile Tyr Leu Lys Phe Gly Thr Asp Gly Gly Ser Ser Tyr Met		
2315	2320	2325
Gly Phe Lys Ala Lys Tyr Ser Ile Ala Ser Cys Gly Gly Thr Val		
2330	2335	2340
Ser Gly Asp Ser Gly Val Ile Glu Ser Ile Gly Tyr Pro Thr Leu		
2345	2350	2355
Pro Tyr Ala Asn Asn Val Phe Cys Gln Trp Phe Ile Arg Gly Leu		
2360	2365	2370
Pro Gly Gly Tyr Leu Thr Leu Ser Phe Glu Asp Phe Asn Leu Gln		

2375	2380	2385
Ser Ser Pro Gly Cys Thr Lys Asp Phe Val Glu Ile Trp Glu Asn		
2390	2395	2400
Gly Thr Ser Gly Arg Val Leu Gly Arg Tyr Cys Gly Asn Ser Thr		
2405	2410	2415
Pro Ser Ser Val Asp Thr Ser Ser Asn Val Ala Ser Val Lys Phe		
2420	2425	2430
Val Thr Asp Gly Ser Val Thr Ala Ser Gly Phe Arg Leu Gln Phe		
2435	2440	2445
Lys Ser Ser Arg Gln Val Cys Gly Gly Asp Leu Gly Gly Pro Thr		
2450	2455	2460
Gly Thr Phe Thr Ser Pro Asn Tyr Pro Asn Pro Asn Pro Gly Ala		
2465	2470	2475
Arg Ile Cys Glu Trp Thr Ile Thr Val Gln Glu Gly Arg Arg Ile		
2480	2485	2490
Val Leu Thr Phe Thr Asn Leu Arg Leu Ser Thr Gln Pro Ser Cys		
2495	2500	2505
Asn Ser Glu Gly Leu Ile Val Phe Asn Gly Ile Arg Ser Asn Ser		
2510	2515	2520
Pro Leu Leu Gln Lys Leu Cys Ser Arg Val Asn Val Thr Asn Glu		
2525	2530	2535
Phe Lys Ser Ser Gly Asn Thr Met Lys Val Val Phe Phe Thr Asp		
2540	2545	2550
Gly Ser Arg Pro Tyr Gly Gly Phe Thr Ala Ser Tyr Thr Ser Thr		
2555	2560	2565
Glu Asp Ala Val Cys Gly Gly Phe Leu Pro Ser Val Ser Gly Gly		
2570	2575	2580
Asn Phe Ser Ser Pro Gly Tyr Asn Gly Ile Arg Asp Tyr Ala Arg		
2585	2590	2595
Asn Leu Asp Cys Glu Trp Thr Leu Ser Asn Pro Asn Arg Glu Asn		
2600	2605	2610
Ser Ser Ile Ser Ile Tyr Phe Leu Glu Leu Ser Ile Glu Ser Gly		
2615	2620	2625
Gln Asp Cys Thr Phe Asp Val Leu Glu Phe Arg Val Gly Asp Ala		
2630	2635	2640
Asp Gly Pro Leu Ile Glu Lys Phe Cys Ser Leu Ser Ala Pro Thr		
2645	2650	2655
Ala Pro Leu Val Ile Pro Tyr Pro Gln Val Trp Ile Arg Phe Val		

2660	2665	2670
Ser Asn Glu Arg Val Glu Tyr Thr Gly Phe Tyr Ile Glu Tyr Ser		
2675	2680	2685
Phe Thr Asp Cys Gly Gly Ile Arg Thr Gly Asp Asn Gly Val Ile		
2690	2695	2700
Ser Ser Pro Asn Tyr Pro Asn Leu Tyr Ser Ala Trp Thr Gly Cys		
2705	2710	2715
Ser Trp Leu Leu Lys Ala Pro Glu Gly Gly Thr Ile Thr Leu Thr		
2720	2725	2730
Leu Ser Asp Phe Leu Leu Glu Ala Gly Pro Thr Cys Thr Ser Asp		
2735	2740	2745
Ser Val Thr Val Arg Asn Gly Asp Ser Pro Gly Ser Pro Val Ile		
2750	2755	2760
Gly Arg Tyr Cys Gly Gln Ser Val Pro Arg Pro Ile Gln Ser Gly		
2765	2770	2775
Ser Asn Gln Leu Ile Val Thr Phe Asn Thr Asn Asn Gln Gly Gln		
2780	2785	2790
Thr Arg Gly Phe Tyr Ala Thr Trp Thr Thr Asn Ala Leu Gly Cys		
2795	2800	2805
Gly Gly Thr Phe Gly Ser Ala Asn Gly Thr Ile Lys Ser Pro Gly		
2810	2815	2820
Trp Pro Gln Thr Phe Pro Glu Asn Ser Arg Cys Ser Trp Thr Val		
2825	2830	2835
Ile Thr Gly Asp Ser Lys Gly Trp Glu Ile Ser Phe Asp Ser Asn		
2840	2845	2850
Phe Arg Ile Pro Ser Ser Asp Ser Gln Cys Gln Asn Ser Phe Val		
2855	2860	2865
Lys Val Trp Gly Gly Arg Leu Met Ile Asn Lys Thr Leu Leu Ala		
2870	2875	2880
Thr Ser Cys Gly Asp Val Ala Pro Ser Pro Ile Val Thr Ser Gly		
2885	2890	2895
Asn Ile Phe Thr Ala Val Phe Gln Ser Glu Glu Met Ala Ala Gln		
2900	2905	2910
Gly Phe Ser Ala Ser Phe Ile Ser Arg Cys Gly Arg Thr Phe Asn		
2915	2920	2925
Thr Ser Pro Gly Asp Ile Ile Ser Pro Asn Phe Pro Lys Gln Tyr		
2930	2935	2940
Asp Asn Asn Met Asn Cys Thr Tyr Leu Ile Asp Ala Asp Pro Gln		

2945	2950	2955
Ser Leu Val Ile Leu Thr Phe Val Ser Phe Gly Leu Glu Asp Arg		
2960	2965	2970
Ser Ala Ile Thr Gly Thr Cys Asp Gly Asp Gly Leu Gly Ile Ile		
2975	2980	2985
Lys Gly Arg Asn Leu Ser Ser Thr Pro Leu Val Thr Ile Cys Gly		
2990	2995	3000
Ser Glu Thr Leu Arg Pro Leu Thr Val Asp Gly Pro Val Leu Leu		
3005	3010	3015
Asn Phe Tyr Ser Asp Ala Tyr Thr Thr Asp Phe Gly Phe Lys Ile		
3020	3025	3030
Ser Tyr Arg Ala Ile Thr Cys Gly Gly Ile Tyr Asn Glu Ser Ser		
3035	3040	3045
Gly Ile Leu Arg Ser Pro Ser Tyr Ser Tyr Ser Asn Tyr Pro Asn		
3050	3055	3060
Asn Leu Tyr Cys Val Tyr Ser Leu Gly Val Arg Ser Ser Arg Val		
3065	3070	3075
Ile Ile Ile Arg Phe Asn Asp Phe Asp Val Ala Pro Ser Asn Leu		
3080	3085	3090
Cys Ala Gly Asp Phe Leu Glu Val Phe Asp Gly Pro Ser Ile Gly		
3095	3100	3105
Asn Arg Ser Leu Gly Lys Phe Cys Gly Ser Thr Arg Pro Gln Thr		
3110	3115	3120
Val Lys Ser Thr Asn Ser Ser Leu Thr Leu Leu Phe Lys Thr Asp		
3125	3130	3135
Ser Ser Gln Thr Ala Arg Gly Trp Lys Ile Phe Phe Arg Glu Thr		
3140	3145	3150
Ile Gly Pro Gln Gln Gly Cys Gly Gly Tyr Leu Thr Glu Asp Asn		
3155	3160	3165
Gln Ser Phe Val Ser Pro Asp Ser Asp Ser Asn Gly Arg Tyr Asp		
3170	3175	3180
Lys Gly Leu Ser Cys Ile Trp Tyr Ile Val Ala Pro Glu Asn Lys		
3185	3190	3195
Leu Val Lys Leu Thr Phe Asn Val Phe Thr Leu Glu Gly Pro Ser		
3200	3205	3210
Ser Ala Gly Ser Cys Val Tyr Asp Tyr Val Gln Ile Ala Asp Gly		
3215	3220	3225
Ala Ser Ile Asn Ser Tyr Leu Gly Gly Lys Phe Cys Gly Ser Arg		

3230	3235	3240
Met Pro Ala Pro Phe Ile Ser Ser Gly Tyr Phe Leu Thr Phe Gln		
3245	3250	3255
Phe Val Ser Asp Val Thr Val Glu Met Arg Gly Phe Asn Ala Thr		
3260	3265	3270
Tyr Thr Phe Val Asp Met Pro Cys Gly Gly Thr Tyr Asn Ala Thr		
3275	3280	3285
Ser Thr Pro Gln Asn Ala Ser Ser Pro Gly Leu Ser Asn Ile Gly		
3290	3295	3300
Arg Pro Tyr Ser Thr Cys Thr Trp Val Ile Ala Ala Pro Pro Gln		
3305	3310	3315
Gln Gln Val Gln Ile Thr Val Trp Asp Leu Gln Leu Pro Ser Gln		
3320	3325	3330
Asp Cys Ser Gln Ser Tyr Leu Glu Leu Gln Asp Ser Val Gln Thr		
3335	3340	3345
Gly Gly Asn Arg Val Thr Gln Phe Cys Gly Ala Asn Tyr Thr Thr		
3350	3355	3360
Leu Pro Val Phe Tyr Ser Ser Met Ser Thr Ala Val Val Val Phe		
3365	3370	3375
Lys Ser Gly Val Ile Asn Arg Asn Ser Gln Val Gln Phe Ser Tyr		
3380	3385	3390
Gln Ile Ala Asp Cys Asn Arg Glu Tyr Asn Gln Thr Phe Gly Asn		
3395	3400	3405
Leu Lys Ser Pro Gly Trp Pro Gln Asn Tyr Asp Asn Asn Leu Asp		
3410	3415	3420
Cys Thr Ile Ile Leu Arg Ala Pro Gln Asn Gly Ser Ile Ser Leu		
3425	3430	3435
Phe Phe Tyr Trp Phe Gln Leu Glu Asp Ser Arg Gln Cys Met Asn		
3440	3445	3450
Asp Phe Leu Glu Val Arg Asn Gly Gly Ser Ser Thr Ser Pro Leu		
3455	3460	3465
Leu Asp Lys Tyr Cys Ser Asn Leu Leu Pro Asn Pro Val Phe Ser		
3470	3475	3480
Gln Ser Asn Glu Leu Tyr Leu Gly Phe Gly Ser Asp Gly Ser Val		
3485	3490	3495
Thr Asn Asn Gly Tyr Glu Ile Ile Trp Thr Ser Ser Ala Ala Gly		
3500	3505	3510
Cys Gly Gly Thr Leu Leu Gly Asp Glu Gly Ile Phe Thr Asn Pro		

3515	3520	3525
Gly Phe Pro Asp Ser Tyr Pro Asn Asn Thr Gly Cys Glu Trp Thr		
3530	3535	3540
Ile Val Ala Pro Ser Gly Arg Pro Val Ser Val Gly Phe Pro Phe		
3545	3550	3555
Leu Ser Ile Asp Ser Ser Gly Gly Cys Asp Gln Asn Tyr Leu Ile		
3560	3565	3570
Val Phe Asn Gly Pro Asp Ala Asn Ser Pro Pro Phe Gly Pro Leu		
3575	3580	3585
Cys Gly Ile Asn Thr Gly Ile Ala Pro Phe Tyr Ala Ser Ser Asn		
3590	3595	3600
Arg Val Phe Ile Arg Phe Gly Ala Glu Tyr Thr Thr Arg Leu Ser		
3605	3610	3615
Gly Phe Glu Ile Met Trp Ser Ser		
3620		

<210>	3
<211>	22
<212>	DNA
<213>	artificial sequence
<220>	
<221>	primer_bind
<222>	bps 838-859 of rat <i>cubilin</i>
<223>	primer for 5' RACE
<400>	3

acacaaggct ccttctactg tg	22
--------------------------	----

<210>	4
<211>	20
<212>	DNA
<213>	artificial sequence
<220>	
<221>	primer_bind
<222>	bps 6872-6891 of rat <i>cubilin</i>
<223>	primer for 3' RACE
<400>	4

gtctggtttc caagttgtgt	20
-----------------------	----

<210> 5
<211> 21
<212> DNA
<213> artificial sequence
<220>
<221> primer_bind
<222> bps 7152-7172 of rat cubilin
<223> primer for 3' RACE
<400> 5

tcagagctct cctggttgtta c 21

<210> 6
<211> 36
<212> PRT
<213> rat
<220>
<223> amino acid sequence of rat cubilin EGF1 repeat
<400> 6

Arg Lys Val Cys Ser Ser Asn Pro Cys Leu Asn Gly Gly Thr Cys
5 10 15
Val Asn Leu His Asp Ser Phe Val Cys Ile Cys Pro Ser Gln Trp
20 25 30
Lys Gly Leu Phe Cys Ser
35

<210> 7
<211> 43
<212> PRT
<213> rat
<220>
<223> amino acid sequence of rat cubilin EGF2 repeat
<400> 7

Glu Asp Val Asn Glu Cys Val Val Tyr Ser Gly Thr Pro Phe Gly
5 10 15
Cys Gln Ser Gly Ser Thr Cys Val Asn Thr Val Gly Ser Phe Arg
20 25 30

Cys Asp Cys Thr Pro Asp Thr Tyr Gly Pro Gln Cys Ala

35 40

<210> 8
<211> 50
<212> PRT
<213> rat
<220>
<223> amino acid sequence of rat cubilin EGF3 repeat
<400> 8

Ser Lys Tyr Asn Asp Cys Glu Gln Gly Ser Lys Gln Leu Cys Lys
5 10 15

His Gly Ile Cys Glu Asp Leu Gln Arg Val His His Gly Gln Pro
20 25 30

Asn Phe His Cys Ile Cys Asp Ala Gly Trp Thr Thr Pro Pro Asn
35 40 45

Gly Ile Ser Cys Thr
50

<210> 9
<211> 43
<212> PRT
<213> rat
<220>
<223> amino acid sequence of rat cubilin EGF4 repeat
<400> 9

Glu Asp Lys Asp Glu Cys Ser Leu Gln Pro Ser Pro Cys Ser Glu
5 10 15

His Ala Gln Cys Phe Asn Thr Gln Gly Ser Phe Tyr Cys Gly Ala
20 25 30

Cys Pro Lys Gly Trp Gln Gly Asn Gly Tyr Glu Cys Gln

35 40

<210> 10
<211> 44
<212> PRT
<213> rat

<220>

<223> amino acid sequence of rat cubilin EGF5 repeat

<400> 10

Asp Ile Asn Lys Cys Glu Ile Asn Asn Gly Gly Cys Ser Gln Ala

5 10 15

Pro Leu Val Pro Cys Leu Asn Thr Pro Gly Ser Phe Ser Cys Gly

20 25 30

Asn Cys Pro Ala Gly Phe Ser Gly Asp Gly Arg Val Cys Thr

35 40

<210> 11

<211> 48

<212> PRT

<213> rat

<220>

<223> amino acid sequence of rat cubilin EGF6 repeat

<400> 11

Pro Val Asp Ile Cys Ser Ile His Asn Gly Gly Cys His Pro Glu

5 10 15

Ala Thr Cys Ser Ser Ser Pro Val Leu Gly Ser Phe Leu Pro Val

20 25 30

Cys Thr Cys Pro Pro Gly Tyr Thr Gly Asn Gly Tyr Gly Ser Asn

35 40 45

Gly Cys Val

<210> 12

<211> 37

<212> PRT

<213> rat

<220>

<223> amino acid sequence of rat cubilin EGF7 repeat

<400> 12

Arg Leu Ser Asn Ile Cys Ser Arg His Pro Cys Val Asn Gly Gln

5 10 15

Cys Ile Glu Thr Val Ser Ser Tyr Phe Cys Lys Cys Asp Ser Gly

20 25 30

Trp Ser Gly Gln Asn Cys Thr

35

<210> 13
 <211> 38
 <212> PRT
 <213> rat
 <220>
 <223> amino acid sequence of rat cubilin EGF8 repeat
 <400> 13

Glu Asn Ile Asn Asp Cys Ser Ser Asn Pro Cys Leu Asn Gly Gly
 5 10 15

Thr Cys Ile Asp Gly Ile Asn Gly Phe Thr Cys Asp Cys Thr Ser
 20 25 30

Ser Trp Thr Gly Tyr Tyr Cys Gln

35

<210> 14
 <211> 42
 <212> PRT
 <213> human
 <220>
 <223> amino acid sequence of Bmp-1 EGF1 repeat
 <400> 14

Glu Val Asp Glu Cys Ser Arg Pro Asn Arg Gly Gly Cys Glu Gln
 5 10 15

Arg Cys Leu Asn Thr Leu Gly Ser Tyr Lys Cys Ser Cys Asp Pro
 20 25 30

Gly Tyr Glu Leu Ala Pro Asp Lys Arg Arg Cys Glu

35 40

<210> 15
 <211> 41
 <212> PRT
 <213> *Drosophila*
 <220>
 <223> amino acid sequence of Tolloid EGF1 repeat

<400> 15

Asp Val Asp Glu Cys Lys Phe Thr Asp His Gly Cys Gln His Leu
5 10 15
Cys Ile Asn Thr Leu Gly Ser Tyr Gln Cys Gly Cys Arg Ala Gly
20 25 30
Tyr Glu Leu Gln Ala Asn Gly Lys Thr Cys Glu
35 40

<210> 16

<211> 41

<212> PRT

<213> *Drosophila*

<220>

<223> amino acid sequence of Tolloid EGF2 repeat

<400> 16

Asp Val Asp Glu Cys Ser Met Asn Asn Gly Gly Cys Gln His Arg
5 10 15
Cys Arg Asn Thr Phe Gly Ser Tyr Gln Cys Ser Cys Arg Asn Gly
20 25 30
Tyr Thr Leu Ala Glu Asn Gly His Asn Cys Thr
35 40

<210> 17

<211> 42

<212> PRT

<213> human

<220>

<223> amino acid sequence of C1s EGF1 repeat

<400> 17

Asp Ile Asn Glu Cys Thr Asp Phe Val Asp Val Pro Cys Ser His
5 10 15
Phe Cys Asn Asn Phe Ile Gly Gly Tyr Phe Cys Ser Cys Pro Pro
20 25 30
Glu Tyr Phe Leu His Asp Asp Met Lys Asn Cys Gly
35 40

<210> 18

<211> 42
<212> PRT
<213> human
<220>
<223> amino acid sequence of fibrillin-1 EGF5 repeat
<400> 18

Asp Ile Asp Glu Cys Ser Thr Ile Pro Gly Ile Cys Glu Gly Gly
5 10 15
Glu Cys Thr Asn Thr Val Ser Ser Tyr Phe Cys Lys Cys Pro Pro
20 25 30
Gly Phe Tyr Thr Ser Pro Asp Gly Thr Arg Cys Ile
35 40

<210> 19
<211> 41
<212> PRT
<213> human
<220>
<223> amino acid sequence of fibrillin-1 EGF13 repeat
<400> 19

Asp Ile Asp Glu Cys Glu Ser Ser Pro Cys Ile Asn Gly Val Cys
5 10 15
Lys Asn Ser Pro Gly Ser Phe Ile Cys Glu Cys Ser Ser Glu Ser
20 25 30
Thr Leu Asp Pro Lys Thr Lys Thr Ile Cys Ile
35 40

<210> 20
<211> 41
<212> PRT
<213> human
<220>
<223> amino acid sequence of fibrillin-1 EGF26 repeat
<400> 20

Asp Val Asn Glu Cys Leu Asp Pro Thr Thr Cys Ile Ser Gly Asn
5 10 15

Cys	Val	Asn	Thr	Pro	Gly	Ser	Tyr	Ile	Cys	Asp	Cys	Pro	Pro	Asp
			20					25					30	
Phe	Glu	Leu	Asn	Pro	Thr	Arg	Val	Gly	Cys	Val				
				35				40						

<210>	21													
<211>	110													
<212>	PRT													
<213>	rat													
<220>														
<223>	amino acid sequence of rat cubilin CUB2 domain													
<400>	21													

Cys	Gly	Gly	Ile	Leu	Thr	Asp	Asn	Tyr	Gly	Ser	Ile	Thr	Ser	Pro
								5		10			15	
Gly	Tyr	Pro	Gly	Asn	Tyr	Pro	Pro	Gly	Arg	Asp	Cys	Val	Trp	Gln
							20			25			30	
Val	Leu	Val	Asn	Pro	Asn	Ser	Leu	Ile	Thr	Phe	Thr	Phe	Gly	Thr
								35			40		45	
Leu	Ser	Leu	Glu	Ser	His	Asn	Asp	Cys	Ser	Lys	Asp	Tyr	Leu	Glu
							50			55			60	
Ile	Arg	Asp	Gly	Pro	Phe	His	Gln	Asp	Pro	Val	Leu	Gly	Lys	Phe
							65			70			75	
Cys	Thr	Ser	Leu	Ser	Thr	Pro	Pro	Leu	Lys	Thr	Thr	Gly	Pro	Ala
							80			85			90	
Ala	Arg	Ile	His	Gly	His	Ser	Cys	Ser	Glu	Thr	Ser	Asp	Lys	Gly
							95			100			105	
Phe	His	Ile	Thr	Tyr										
				110										

<210>	22													
<211>	108													
<212>	PRT													
<213>	rat													
<220>														
<223>	amino acid sequence of rat cubilin CUB5 domain													
<400>	22													

Cys Gly Glu Val Leu Thr Ala Ser Thr Gly Ile Ile Glu Ser Pro
 5 15
 Gly His Pro Asn Val Tyr Pro Arg Gly Val Asn Cys Thr Trp His
 20 30
 Val Val Val Gln Arg Gly Gln Leu Ile Arg Leu Glu Phe Ser Ser
 35 45
 Phe Tyr Leu Glu Phe His Tyr Asn Cys Thr Asn Asp Tyr Leu Glu
 50 60
 Ile Tyr Asp Thr Ala Ala Gln Thr Phe Leu Gly Arg Tyr Cys Gly
 65 75
 Lys Ser Ile Pro Pro Ser Leu Thr Ser Asn Ser Asn Ser Ile Lys
 80 90
 Leu Ile Phe Val Ser Asp Ser Ala Leu Ala His Glu Phe Gly Ser
 95 100 105
 Ile Asn Tyr

<210> 23
 <211> 111
 <212> PRT
 <213> rat
 <220>
 <223> amino acid sequence of rat cubilin CUB6 domain
 <400> 23

Cys Leu Tyr Asp Tyr Thr Asp Asn Phe Gly Met Leu Ser Ser Pro
 5 15
 Asn Phe Pro Asn Asn Tyr Pro Ser Asn Trp Glu Cys Ile Tyr Arg
 20 30
 Ile Thr Val Gly Leu Asn Gln Gln Ile Ala Leu His Phe Thr Asp
 35 45
 Phe Thr Leu Glu Asp Tyr Phe Gly Ser Gln Cys Val Asp Phe Val
 50 60
 Glu Ile Arg Asp Gly Gly Tyr Glu Thr Ser Pro Leu Val Gly Ile
 65 75
 Tyr Cys Gly Ser Val Leu Pro Pro Thr Ile Ile Ser His Ser Asn
 80 90
 Lys Leu Trp Leu Lys Phe Lys Ser Asp Ala Ala Leu Thr Ala Lys
 95 100 105
 Gly Phe Ser Ala Tyr Trp

110

<210> 24
<211> 113
<212> PRT
<213> rat
<220>
<223> amino acid sequence of rat cubilin CUB9 domain
<400> 24

<210> 25
<211> 110
<212> PRT
<213> rat
<220>
<223> amino acid sequence of rat cubilin CUB12 domain
<400> 25

Cys	Gly	Gly	Ser	Phe	Tyr	Thr	Leu	Asp	Gly	Ile	Phe	Asn	Ser	Pro
				5					10					15
Asp	Tyr	Pro	Ala	Asp	Tyr	His	Gly	Asn	Ala	Glu	Cys	Val	Trp	Asn
				20					25					30

Ile Ala Ser Ser Pro Gly Asn Arg Leu Gln Leu Ser Phe Leu Ser
 35 40 45
 Phe Asn Leu Glu Asn Ser Leu Asn Cys Asn Lys Asp Phe Val Glu
 50 55 60
 Ile Arg Glu Gly Asn Ala Thr Gly His Leu Ile Gly Arg Tyr Cys
 65 70 75
 Gly Asn Ser Leu Pro Gly Asn Tyr Ser Ser Ala Glu Gly His Ser
 80 85 90
 Leu Trp Val Arg Phe Val Ser Asp Gly Ser Gly Thr Gly Met Gly
 95 100 105
 Phe Gln Ala Arg Phe
 110

<210> 26
 <211> 110
 <212> PRT
 <213> rat
 <220>
 <223> amino acid sequence of rat cubilin CUB17 domain
 <400> 26

Cys Gly Gly Thr Val Ser Gly Asp Ser Gly Val Ile Glu Ser Ile
 5 10 15
 Gly Tyr Pro Thr Leu Pro Tyr Ala Asn Asn Val Phe Cys Gln Trp
 20 25 30
 Phe Ile Arg Gly Leu Pro Gly His Tyr Leu Thr Leu Ser Phe Glu
 35 40 45
 Asp Phe Asn Leu Gln Ser Ser Pro Gly Cys Thr Lys Asp Phe Val
 50 55 60
 Glu Ile Trp Glu Asn His Thr Ser Gly Arg Val Leu Gly Arg Tyr
 65 70 75
 Cys Gly Asn Ser Thr Pro Ser Ser Val Asp Thr Ser Ser Asn Val
 80 85 90
 Ala Ser Val Lys Arg Val Thr Asp Gly Ser Val Thr Ala Ser Gly
 95 100 105
 Phe Arg Leu Gln Phe
 110

<210> 27
<211> 110
<212> PRT
<213> rat
<220>
<223> amino acid sequence of rat cubilin CUB20 domain
<400> 27

<210> 28
<211> 110
<212> PRT
<213> human
<220>
<223> amino acid sequence of Bmp-1 CUB1 domain
<400> 28

Cys	Gly	Glu	Thr	Leu	Gln	Asp	Ser	Thr	Gly	Asn	Phe	Ser	Ser	Pro
								5						15
Glu	Tyr	Pro	Asn	Gly	Tyr	Ser	Ala	His	Met	His	Cys	Val	Trp	Arg
								20						30
Ile	Ser	Val	Thr	Pro	Gly	Glu	Lys	Ile	Ile	Leu	Asn	Phe	Thr	Ser
								35						45

Leu Asp Leu Tyr Arg Ser Arg Leu Cys Trp Tyr Asp Tyr Val Glu
 50 55 60
 Val Arg Asp Gly Phe Trp Arg Lys Ala Pro Leu Arg Gly Arg Phe
 65 70 75
 Cys Gly Ser Lys Leu Pro Glu Pro Ile Val Ser Thr Asp Ser Arg
 80 85 90
 Leu Trp Val Glu Phe Arg Ser Ser Asn Trp Val Gly Lys Gly
 95 100 105
 Phe Phe Ala Val Tyr
 110

<210> 29
 <211> 110
 <212> PRT
 <213> human
 <220>
 <223> amino acid sequence of Bmp-1 CUB2 domain
 <400> 29

Cys Gly Gly Asp Val Lys Lys Asp Tyr Gly His Ile Gln Ser Pro
 5 10 15
 Asn Tyr Pro Asp Asp Tyr Arg Pro Ser Lys Val Cys Ile Trp Arg
 20 25 30
 Ile Gln Val Ser Glu Gly Phe His Val Gly Leu Thr Phe Gln Ser
 35 40 45
 Phe Glu Ile Glu Arg His Asp Ser Cys Ala Tyr Asp Tyr Leu Glu
 50 55 60
 Val Arg Asp Gly His Ser Glu Ser Ser Thr Leu Ile Gly Arg Tyr
 65 70 75
 Cys Gly Tyr Glu Lys Pro Asp Asp Ile Lys Ser Thr Ser Ser Arg
 80 85 90
 Leu Trp Leu Lys Phe Val Ser Asp Gly Ser Ile Asn Lys Ala Gly
 95 100 105
 Phe Ala Val Asn Phe
 110

<210> 30
 <211> 111
 <212> PRT

<213> *Drosophila*

<220>

<223> amino acid sequence of Tollid CUB2 domain

<400>

<210> 31

<211> 116

<212> PBT

<213> *Prospephila*

<220>

<223> amino acid sequence of Tolloid CUB3 domain
<400> 31

...ly Gly Val Val Asp Ala Thr Lys Ser Asn Gly Ser Leu Ty
5 10 15

Ser Pro Ser Tyr Pro Asp Val Tyr Pro Asn Ser Lys Gln Cys Val
20 25 30

Trp Glu Val Val Ala Pro Pro Asn His Ala Val Phe Leu Asn Phe
35 40 45

Ser His Phe Asp Leu Glu Gly Thr Arg Phe His Tyr Thr Lys Cys
 50 55 60

Asn	Tyr	Asp	Tyr	Leu	Ile	Ile	Tyr	Ser	Lys	Met	Arg	Asp	Asn	Arg
65										70				75
Leu	Lys	Lys	Ile	Gly	Ile	Tyr	Cys	Gly	His	Glu	Leu	Pro	Pro	Val
										80			85	90
Val	Asn	Ser	Glu	Gln	Ser	Ile	Leu	Arg	Leu	Glu	Phe	Tyr	Ser	Asp
										95			100	105
Arg	Thr	Val	Gln	Arg	Ser	Gly	Phe	Val	Lys	Phe				
									110				115	

<210>	32													
<211>	109													
<212>	PRT													
<213>	<i>Drosophila</i>													
<220>														
<223>	amino acid sequence of Tolloid CUB4 domain													
<400>	32													

Cys	Lys	Phe	Glu	Ile	Thr	Thr	Ser	Tyr	Gly	Val	Leu	Gln	Ser	Pro
										5				15
Asn	Tyr	Pro	Glu	Asp	Tyr	Pro	Arg	Asn	Ile	Tyr	Cys	Tyr	Trp	His
										20				30
Phe	Gln	Thr	Val	Leu	Gly	Phe	Ile	Gln	Leu	Thr	Phe	His	Asp	Phe
										35				45
Glu	Val	Glu	Ser	His	Gln	Glu	Cys	Ile	Tyr	Asp	Tyr	Val	Ala	Ile
										50				60
Tyr	Asp	Gly	Arg	Ser	Glu	Asn	Ser	Ser	Thr	Leu	Gly	Ile	Tyr	Cys
										65				75
Gly	Gly	Arg	Glu	Pro	Tyr	Ala	Val	Ile	Ala	Ser	Thr	Asn	Glu	Met
										80				90
Phe	Met	Val	Leu	Ala	Thr	Asp	Ala	Gly	Leu	Gln	Arg	Lys	Gly	Phe
										95				105
Lys	Ala	Thr	Phe											

<210>	33													
<211>	109													
<212>	PRT													
<213>	<i>Xenopus laevis</i>													

<220>

<223> amino acid sequence of Uvs-2 CUB2 domain

<400> 33

Cys	Gly	Gly	Ala	Phe	Tyr	Ser	Ser	Pro	Lys	Thr	Phe	Thr	Ser	Pro
				5					10					15
Asn	Tyr	Pro	Gly	Asn	Tyr	Thr	Thr	Asn	Thr	Asn	Cys	Thr	Trp	Thr
				20					25					30
Ile	Thr	Ala	Pro	Ala	Gly	Phe	Lys	Val	Ser	Leu	Arg	Ile	Thr	Asp
				35					40					45
Phe	Glu	Leu	Glu	Ile	Gly	Ala	Ser	Cys	Arg	Tyr	Asp	Tyr	Leu	Asn
				50					55					60
Ile	Tyr	Asn	Ser	Thr	Leu	Gly	Ala	Val	Met	Gly	Pro	Tyr	Cys	Gly
				65					70					75
Pro	Ile	Asp	Phe	His	Ser	Ala	Ile	Val	Ser	Lys	Ser	Asn	Ser	Met
				80					85					90
Met	Ile	Thr	Met	Asn	Ser	Asp	Phe	Ser	Lys	Gln	Tyr	Lys	Gly	Phe
				95					100					105
Ser	Ala	Thr	Tyr											

<210> 34

<211> 112

<212> PRT

<213> human

<220>

<223> amino acid sequence of C1s CUB1 domain

<400> 34

Glu	Pro	Thr	Met	Tyr	Gly	Glu	Ile	Leu	Ser	Pro	Asn	Tyr	Pro	Gln
				5					10					15
Ala	Tyr	Pro	Ser	Glu	Val	Glu	Lys	Ser	Trp	Asp	Ile	Glu	Val	Pro
				20					25					30
Glu	Gly	Tyr	Gly	Ile	His	Leu	Tyr	Phe	Thr	His	Leu	Asp	Ile	Glu
				35					40					45
Leu	Ser	Glu	Asn	Cys	Ala	Tyr	Asp	Ser	Val	Gln	Ile	Ile	Ser	Gly
				50					55					60
Asp	Thr	Glu	Glu	Gly	Arg	Leu	Cys	Gly	Gln	Arg	Ser	Ser	Asn	Asn
				65					70					75

Pro His Ser Pro Ile Val Glu Glu Phe Gln Val Pro Tyr Asn Lys		
80	85	90
Leu Gln Val Ile Phe Lys Ser Asp Phe Ser Asn Glu Glu Arg Phe		
95	100	105
Thr Gly Phe Ala Ala Tyr Tyr		
110		

<210>	35	
<211>	109	
<212>	PRT	
<213>	human	
<220>		
<223>	amino acid sequence of Tsg6 CUB domain	
<400>	35	

Cys Gly Gly Val Gly Thr Asp Pro Lys Arg Ile Phe Lys Ser Pro		
5	10	15
Gly Phe Pro Asn Glu Tyr Glu Asp Asn Gln Ile Cys Tyr Trp His		
20	25	30
Ile Arg Leu Lys Tyr Gly Gln Arg Ile His Leu Ser Phe Leu Asp		
35	40	45
Phe Asp Leu Glu Asp Asp Pro Gly Cys Leu Ala Asp Tyr Val Glu		
50	55	60
Ile Tyr Asp Ser Tyr Asp Asp Val His Gly Phe Val Gly Arg Tyr		
65	70	75
Cys Gly Asp Glu Leu Pro Asp Asp Ile Ile Ser Gly Asn Val Met		
80	85	90
Thr Leu Lys Phe Leu Ser Asp Ala Ser Val Thr Ala Gly Gly Phe		
95	100	105
Gln Ile Lys Tyr		

<210>	36	
<211>	98	
<212>	PRT	
<213>	pig	
<220>		
<223>	amino acid sequence of Aqn-3 CUB domain	

<400> 36

Cys Gly Gly Phe Leu Lys Asn Tyr Ser Gly Trp Ile Ser Tyr Tyr
5 10 15
Lys Ala Leu Thr Thr Asn Cys Val Trp Thr Ile Glu Met Lys Pro
20 25 30
Gly His Lys Ile Ile Leu Gln Ile Leu Pro Leu Asn Leu Thr Cys
35 40 45
Lys Glu Tyr Leu Glu Val Arg Asp Gln Arg Ala Gly Pro Asp Asn
50 55 60
Phe Leu Lys Val Cys Gly Gly Thr Gly Phe Val Tyr Gln Ser Ser
65 70 75
His Asn Val Ala Thr Val Lys Tyr Ser Arg Asp Ser His His Pro
80 85 90
Ala Ser Ser Phe Asn Val Tyr Phe
95

<210> 37

<211> 21

<212> DNA

<213> artificial sequence

<220>

<221> primer_bind

<223> a cubilin primer used for RT-PCR

<400> 37

tgcctaccac agcccaaatg a 21

<210> 38

<211> 19

<212> DNA

<213> artificial sequence

<220>

<221> primer_bind

<223> a cubilin primer used for RT-PCR

<400> 38

agagccacaa tgactgcag 19

<210> 39

<211> 22
<212> DNA
<213> artificial sequence
<220>
<221> primer_bind
<223> a megalin primer used for RT-PCR
<400> 39

gccaggaga caggaacagt ag

22

<210> 40
<211> 22
<212> DNA
<213> artificial sequence
<220>
<221> primer_bind
<223> a megalin primer used for RT-PCR
<400> 40

tcacaaaaatg ccagaccacg aa

22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/01259

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/00, 15/12; C07K 14/705, 14/435
 US CL : 435/69.1, 320.1, 325, 6; 536/23.5, 23.1; 530/350

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 320.1, 325, 6; 536/23.5, 23.1; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	MOESTRUP, S.K. et al. The intrinsic factor-vitamin B ₁₂ receptor and target of teratogenic antibodies is a megalin-binding peripheral membrane protein with homology to developmental proteins. The Journal of Biological Chemistry. 27 February 1998, Vol. 273, No. 9, pages 5235-5242, especially Figures 1 and 2 and Genbank Accession Number AF022247.	1-13
X,P	KOZYRAKI, R. et al. The human intrinsic factor-vitamin B ₁₂ receptor, cubilin: Molecular characterization and chromosomal mapping of the gene to 10p within the autosomal recessive megaloblastic anemia (MGA1) region. Blood. 15 May 1998, Vol. 91, No. 10, pages 3593-3600, especially Figure 3 and attached Genbank Accession Number AF034611.	1-13

 Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	*T*	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance		
B earlier document published on or after the international filing date	*X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
L document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
O document referring to an oral disclosure, use, exhibition or other means	*Z*	document member of the same patent family
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
19 APRIL 1999

Date of mailing of the international search report

14 MAY 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230Authorized officer
ELIANE LAZAR-WESLEY
Telephone No. (703) 308-0196JOYCE BRIDGERS
PARALEGAL SPECIALIST
CHEMICAL MATRIX
Job

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01259

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BIRN, H. et al. Characterization of an epithelial 460-kDa protein that facilitates endocytosis of intrinsic factor-vitamin B ₁₂ and binds receptor-associated protein. The Journal of Biological Chemistry. 17 October 1997, Vol. 272, No. 42, pages 26497-26504, especially page 26499, results, and Figure 6.	1, 4, 10, 11, 13

Form PCT/ISA/210 (continuation of second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01259

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-13

Remark on Protest

The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01259

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, MEDLINE, CAPLUS

search terms: cubilin, megalin, kidney, renal, receptor

GENBANK, EMBL, SWISSPROT

search: SEQ ID No: 1 and 2

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-13, drawn to DNA encoding a cubulin protein, isolated cubulin protein and a method of detecting expression of the protein.

Group II, claims 14-15, drawn to a method of treating or reducing toxicity.

Group III, claims 16-19, drawn to a receptor for ligands.

Group IV, claim 20, drawn to a method of detecting renal damage.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical features of Group I are the DNA encoding cubulin, the cubulin protein and a method of detecting expression of the protein. Groups II-IV do not share the special technical feature of Group I, because the receptor of Group III is a product different of the product of Group I by its structural and functional characteristics, and the methods of Groups II and IV differ from the method of Group I because they use different reagents and different method steps in order to reach different goals.

THIS PAGE BLANK (USPTO)